Essays in the Economics of Migration

Lars Ludolph

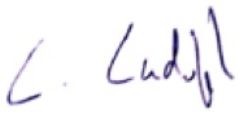
A thesis presented for the degree of
Doctor of Philosophy

Department of Geography and Environment
London School of Economics and Political Science
August 2021

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it). The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 75,000 words, excluding all appendices and the bibliography.



Lars Ludolph

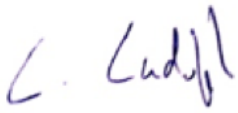
London, 06/08/2021

Statement of conjoint work

The paper "Barriers to humanitarian migration, victimisation and integration outcomes: Evidence from Germany" is co-authored with Teresa Freitas-Monteiro. My contribution amounts to at least 50% of the total work.

The paper "Global prices and migration in Sub-Sahara Africa" is co-authored with Barbora Sedova. My contribution amounts to at least 50% of the total work.

The paper "Immigration and redistributive spending: Evidence from English local authorities" is co-authored with Nikolaj Broberg. My contribution amounts to at least 50% of the total work.



Lars Ludolph

London, 06/08/2021

Acknowledgements

One of the curious thoughts I kept having throughout the PhD is that I am not only researching migration, but live, work and study in an environment where everybody, including myself, is a migrant. I am deeply grateful for my time at the London School of Economics that allowed me to meet and learn from so many people from around the world.

First and foremost, I am indebted to Riccardo Crescenzi and Nancy Holman for their outstanding supervision. Riccardo's ability to see the bigger picture in research projects and Nancy's continuous support and reassuring presence helped my progress enormously. Their guidance was invaluable. I am particularly grateful to Olmo Silva whose ability to spot weaknesses in empirical strategies has sent me back to the drawing board a few times. My special thank you also goes to Mikkel Barslund whose early-stage support made this PhD project possible. Our discussions about every chapter of this project were always inspiring.

Doing research is both more productive and fun when doing it in a team. I would like to thank my co-authors, Nikolaj Broberg, Teresea Freitas-Monteiro and Barbora Sedova for all the productive exchanges. I learnt a lot from you and thoroughly enjoyed our cooperation.

A particularly valuable asset at the department is the Economic Geography work-in-progress seminar series. I would like to mention Gabriel Ahlfeldt, Felipe Carozzi, Steve Gibbons, Christian Hilber, Neil Lee, Henry Overman, Lindsay Relihan and Simona Iammarino among many others for their valuable comments during my presentations.

My work benefited from the feedback received during workshop and seminar presentations at the European University Institute, the German Institute for Economic Research, Glasgow Caledonian University, the Ifo institute for economic research, the German Institute for employment research, the Maastricht School of Governance, Maastricht University, the Max Planck Institute, the Migration Policy Centre, Oxford University, and the Economics of Migration Junior Online Seminar organised by LISER, PSE and UC3M. Chapter 2 further benefited from comments and suggestions by Achim Ahrens, Herbert Brücker, Timo Hener, Markus Nagler, Yuliya Kosyakova and Angela Kunzler. Mikkel

Barslund, Matthew Sharp and Olmo Silva provided valuable comments on chapter 3. I would also like to thank the Austrian Social Science Data Archive and Statistics Austria for their support in facilitating my access to Austrian microcensus data, and Iris Butzlaff in particular, who went above and beyond to make my research possible. Chapter 4 received valuable feedback from Joyce J. Chen and Wolfram Schlenker. Finally, chapter 5 has benefited from in-depth conversations with Brian Bell, Elias Dinas, Jérôme Gonnot, Andrea Ichino, Andrea Mattozzi, Mauro Lanati, David Phillips, Filipa Sá and Danny Walker.

Before the Covid-19 pandemic disrupted our lives, sharing a PhD office with fellow PhD students was a great source of inspiration. Thank you to Andreas, Eduardo, Lindsey, Louise, Margarida, Matt, Sandro, Yohan and many others for the great fun during countless lunch and coffee breaks.

Finally, and above all, I would like to thank my partner, Veselina, who has been the best support I could have asked for. Without you, embarking on this journey would have not been possible. This work is dedicated to you, even though you would have certainly preferred a novel.

Abstract

This thesis comprises four empirical essays that examine different issues in the economics of migration. The common theme to all four essays is the idea that migration is a phenomenon with economic implications for the country of origin, migrants themselves and the destination country. To reap the benefits of migration, it is necessary to understand the challenges and barriers at the various stages of the migratory journey that could reduce welfare and that may require policy interventions - or changes in policy - to overcome them. The research carried out within this PhD project aims to make a contribution to our understanding of the economic aspects of these challenges.

The first essay examines a stage of the humanitarian migration process from developing to developed countries that has thus far been underexplored in economics: The journey itself. Asylum seekers migrating from developing countries to Europe frequently experience victimisation events during their journey. The essay links these potentially traumatic events to economic integration outcomes in Germany, one of the main recipient countries of asylum seekers during the 2015 migration crisis. The study shows that physical victimisation during the journey to Germany is strongly associated with significantly lower mental well-being upon arrival in the destination. The effect on the victimised also leads to a "loss of future directedness", which distorts one of the major decisions newly-arrived migrants have to make: Compared to non-victimised migrants, physically victimised refugees are more likely to engage in part-time and marginal employment instead of pursuing host-country education in the first years after arrival.

The second essay follows up on these findings. The essay analyses the long-term value of formal host-country education for refugees vis-à-vis those the same level of education attained in the country of origin. The study deploys 22 years of Austrian microcensus data and analyses the labour market position of forcibly displaced young Bosnians who arrived in Austria during the 1992-1995 Bosnian war. Exploiting the age at the time of forced migration as an instrument for the probability of receiving host-country instead of origin country education, the results show that attaining a formal degree in the host-country significantly reduces the probability of working in low-quality jobs even after more than two decades of stay in the hosting country.

The third essay shifts the focus to the early stage of the migration process: It studies

the role of household income in developing countries in the decision to send a household member as a labor migrant. The essay analyses the effect of exogenous global crop price changes on migration from agricultural households and finds that migration rates from very poor households indeed increase when the world market price of locally-grown crop rises. The finding suggests that for these households, additional income can relax their liquidity constraint and facilitate migration.

The fourth and final essay then turns the attention back to the destination country. The study analyses the impact of the large, unexpected and spatially heterogenous migration wave from Central and Eastern European countries following their EU accession in 2004 on local level redistributive spending in England. While the arrival of migrants indeed affected public spending and locally generated revenue, the study finds no evidence that these changes in local service provision are driven by a decrease in the local willingness to redistribute following the arrival of outsiders. Rather, the results suggest that the demographic characteristics of Central and Eastern European migrants, in particular their young age, reduced demand for locally supplied social care services.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The movement of people between places has been increasing at a steady pace over the past decades. The number of migrants in the world grew from 174 million in 2000 to 272 million in 2019, while the share of migrants in the total global population increased from 2.8% to 3.5% [IOM, 2020].

If the whole globe was ruled by a benevolent dictator whose main goal it was to maximise global GDP, it is likely that the creation of a world where all restrictions to migration are eliminated would be one of the ruler's first acts. Clemens [2011] suggests that the removal of all barriers to migration would increase global economic output by "one or two orders of magnitude larger than the gains from dropping all remaining restrictions on international flows of goods and capital" [p. 84]. As such, the increase in migration we observed over the past decades is likely to be welfare increasing on the global level. In reality, of course, the question is who captures these gains, which in turn determines the incentives to remove or impose restrictions to migration.

The next section 1.1 provides a high-level overview of the economics along the migratory journey and places the research conducted within this PhD project into this wider framework. It stresses two fundamental ways to distinguish studies on economic aspects of migration: By defining the relevant agent and by characterising the type of migration flow. Against this background, section 1.2 provides a summary of each essay. Finally, section 1.3 discusses some general lessons learnt over the course of this project.

1.1 Economics along the migratory journey

When studying the economics of migration, it is necessary to first define the perspective the study takes. Fundamentally, three groups of economic agents can be distinguished in this regard. The population in countries of origin, migrating individuals themselves or their households and the population in destination countries.

For the country of origin, emigration necessarily means losing parts of their labour force. Unless individuals are forcibly displaced from their home and migrate as humanitarian migrants, a plausible assumption governments in migrant-sending countries should make is that individuals or households migrate because it is a utility-maximising choice. In most cases, this corresponds to higher earning opportunities away from the origin. Constraining the choice set by imposing emigration restrictions is therefore likely to be unpopular with the electorate and only very few governments around the world currently explicitly impose such restrictions. There are, however, instances in which segments of the population left behind may be made worse off through emigration decisions of others. One example is a potential brain drain, i.e. the emigration of highly-skilled individuals. Whether or not the brain drain phenomenon harms economic development depends on a wide range of parameters and losses tend to be borne by low-educated segments of the population [Docquier and Iftikhar, 2019]. Crucially, these parameters include whether the option to migrate can sufficiently incentivise human capital investment, the magnitude of positive diaspora externalities and public policies that stimulate demand for high-skill labour in the country of origin [Docquier and Rapoport, 2012]. Remittances, i.e. the income earned abroad that emigrants transfer back into their country of origin, may cushion these effects by increasing both consumption and investment in recipient households. Yet, the macro-level effect of remittances on economic development in the country of origin remains inconclusive [Rapoport and Docquier, 2006, Yang, 2011].

If migrants flow from poor countries to rich countries in search for better income opportunities, one could assume that once income levels converge, migration rates will decline. However, while the majority of global migrants indeed originate from developing countries and the majority of migrants indeed reside in developed countries [IOM, 2020], research suggests that the relationship between economic development and emigration rates is not linear but rather follows an inverse u-shape [Clemens, 2014]. In the very early stages of development, a rise in income levels increases migration rates. Both macroeconomic and microeconomic reasons explain the phenomenon. On the macro level, the demographic transition associated with rising income at low level of developments that leads to a larger

pool of potential migrants, diaspora-led pull factors once the first migrants settled abroad, employment loss in agriculture, rising inequality that leads the relatively worse-off to consider migration and a higher willingness to receive highly-skilled migrants in destinations all constitute reasons for the observation. Dao et al. [2018] indeed show that rising education levels and macroeconomic factors such as the diaspora abroad explain most of the increasing international migration to OECD countries when income starts rising in initially poor countries. On the micro level, the key reason is the high up-front costs associated with migration. Many liquidity-constrained households in the poorest developing countries cannot afford to send one of their members as a migrant when income levels are low. Once income starts rising, this budget constraint loosens [Bazzi, 2017]. One of the contributions within this PhD projects falls precisely into this literature: In chapter 4, I show that the increasingly volatile global crop prices are positively associated with household out-migration from the poorest (but not the relatively wealthier) agricultural households in five Sub-Saharan African countries. Members of these households rarely migrate to OECD countries but search for employment opportunities internally or in neighbouring African countries. Migration represents an important means of adaptation to changing climatic conditions for these households.

While individuals migrating for labour reasons constitute the main migrant category, employment opportunities are not the only reason why migrants leave their country of origin. According to the United Nations High Commissioner for Refugees (UNHCR), 82.4 million people were forcibly displaced from their homes in 2020, of which 20.7 million resided outside their countries of origin as recognised refugees. The majority of the forcibly displaced reside in developing countries, for the simple reason that most forced displacement occurs in the developing world and the affected seek refuge close to their homes. Naturally, selection dynamics play a less important role among refugees compared to labour migrants when the two populations are compared as a whole. However, this no longer holds true once we further distinguish humanitarian migration flows by their destination. In fact, recent evidence suggests that those who flee to the developed world - where Europe, and Germany in particular, have recently been the main destination - are positively selected on their human capital [Aksoy and Poutvaara, 2019]. One of the key reasons for the positive selection is that severe conflicts decrease the returns to human capital relatively more for the highly-educated who are then more likely to look for opportunities elsewhere. Another likely reason relates to a simple cost-benefit analysis: Both the costs and the risks asylum seekers have to take on when travelling to Europe are enormous, and are only worth-while for those who expect a sufficiently large return

in their intended destination. In chapter 2, I explicitly address the risk asylum seekers face on their journey to Europe and show that victimisation rates are indeed large: In our sample of surveyed refugees in Germany, we find that 36% report having experienced physical victimisation events such as physical abuse, shipwreck or incarceration on their way to their destination.

Victimisation of asylum seekers during the journey to developed countries is a direct consequence of barriers to migration put in place by destination countries. These can be direct, if physical harm is inflicted on asylum seekers by entry-denying border police, or indirect, if the lack of legal migration channels leads humanitarian migrants to rely on potentially violent escape agents. From the perspective of destination countries, these events reflect the consequences of underlying preferences for refugee protection: Evidence suggests that there is a willingness to protect refugees in developed countries but this willingness has clear limits [Jeannet et al., 2021]. While the widespread preference for limiting humanitarian inflows in the developed world may also have non-economic reasons, economic reasons play an important role for public preferences towards the scale of immigration in general. On the simplest level, the economic concerns follow the logic of seminal work by Borjas [1999], who shows that when migration leads to efficiency gains on the host-country labour market, the arrival of newcomers creates winners and losers, with the latter primarily located at the bottom end of the income distribution.

These canonical models of labour supply can help to guide our thinking on the short-term impact of migration on native employment and wages in the destination, but they neglect important general equilibrium considerations such as skill complementarities, native mobility and an increase in aggregate consumption as margins of adjustment [Peri, 2016]. Ignoring the dynamic nature of adjustment processes and time dimensions in general is likely to make predictions on the economic impact of migration inflows both model-dependent and geography-specific [Longhi et al., 2005]. The acknowledgement that the theoretical channels on how migrants impact on the economy in the destination are complex has motivated a copious body of empirical migration impact assessment studies. These impact assessment studies are not limited to the labour market but encompass many dimensions that have first and second-order effects on native economic welfare. Short and long-run migration impact assessments include studies on all dimensions of the labour market, native consumption patterns, health outcomes, crime and security, social mobility, innovation, trade, tourism and many other areas potentially affected by the inflow of outsiders [Nijkamp et al., 2012]. Impact assessment studies can thus take many forms but all have one common denominator: To determine how economies and societies

in destination countries are impacted by migration, which in turn guides crucial questions on migration policy and immigration restrictions.

In particular in countries with generous welfare systems, one of the questions most heatedly debated in public has been the effect of migrants on government spending. Next to fiscal accounting studies on the direct impact of migration on destination countries' public purse (i.e. their net fiscal contribution), a crucial political economy question has been the sustainability of welfare systems if income is not just redistributed between natives, but also towards newly-arrived outsiders. The initial answer to that question appeared to be that heterogenous societies are likely to prefer lower levels of income redistribution, which could explain differences in the generosity of welfare systems between Europe and the United States [Alesina et al., 2001]. The result of a small negative association between migration and redistribution has since been confirmed mostly in cross-country studies [Alesina et al., 2019]. In chapter 5 of this PhD thesis, I speak directly to this literature and show that when testing the hypothesis on the sub-national level in the setting of English local authorities, changes in local public spending patterns in response to migrant inflows are primarily driven by mechanical changes caused by the distinct demographic characteristics of migrants. I find no evidence that regions more exposed to migration inflows changed their voting patterns towards parties running on less redistributive platforms. These results are reconcilable with those found in cross-country studies but imply that the nexus of migration and preferences for lower levels of redistribution observed across countries may not be driven by direct local exposure to migrants.

Once migrants arrive in their destination and acquire the right to work, the preferences of the native population in the destination country and those of migrants are more aligned with regards to migrants' economic integration. Both migrants and their destination countries have an interest in a successful integration: For migrants, a better integration increases income, while for the destination country it directly improves public finances through both a rise in tax income and a decline in welfare spending. Yet, migrants in general and refugees in particular are significantly less likely to be employed than the native-born in developed countries and once they do find work, earn significantly less [Brell et al., 2020]. The general reasons range from the non-transferability of human capital from developing into developed countries to language barriers and restricted access to public sector employment due to the lack of citizenship [Zwysen, 2019]. The refugee-specific reasons include the uncertainty regarding the duration of stay in the destination country and the potential trauma experienced due to the forced displacement [Brell et al.,

2020]. Both chapters 2 and 3 of this PhD thesis focus on aspects of refugee integration. In chapter 2, I show that potentially traumatic events refugees experience often occur during their journey to Europe and these events indeed affect economic integration outcomes. Importantly, they distort one of the crucial decisions refugees have to make when arriving in their destination: To accept the heavy discount on their education attained in the country of origin or to invest into host-country human capital to get access to higher-income jobs at a later stage [Cortes, 2004]. The results suggest that physical victimisation events lead refugees to discount these future payoffs more heavily, leading them to engage in low-quality employment in relatively larger shares. In chapter 3, I then provide causal evidence on the discount refugees face on their formal education attained abroad. I show that formal host-country education significantly reduces low-quality employment among refugees, an effect visible after more than two decades of stay.

1.2 Summary of the PhD thesis

This thesis comprises four empirical essays that examine issues in the economics of migration. It draws on and contributes to literatures in labour economics, the economics of education, economic geography, environmental economics, development economics and political economy but also aims to incorporate insights from the fields of sociology, psychology and medical science.

On the highest level, the PhD thesis can be divided into two parts. The first part includes two contributions that set out with the objective to add to the knowledge on the factors that shape the economic integration success of refugees in their host countries. The second part of the thesis then focuses and expands on more traditional questions related to labour migration, both from the country of origin perspective and the perspective of receiving countries. The following presents a summary of the individual chapters.

1.2.1 First Essay

The first essay explores a stage of the migratory journey of humanitarian migration that has so not been addressed in the economics literature: The journey itself. The starting point of the paper is the observation that this journey is often precarious for asylum seekers from developing countries who seek refuge in the developed world. It then analyses the implications of victimisation events along the journey for the economic integration of new-arrivals in the destination country using refugee survey data collected in the aftermath of the 2015 refugee crisis in Germany. The key finding of the study is that physical victimisation events negatively affect the human capital investment in the destination country and thus distort one of the key decisions refugees have to make: Whether to accept the heavy discount on their foreign human capital and join the labour force as de-facto low-skilled workers, or to invest into host-country education with the aim to obtain higher quality employment at a later stage. Ruling out all alternative possible mechanisms that could explain the results, I place my findings into sociology and psychology literatures, which have found a "loss of future directedness" among the victimised.

1.2.2 Second Essay

The finding of the first essay motivates the second essay, which explicitly analyses the long-term value of formal host-country human capital for refugees. The paper starts with

the observation that refugees hosted across the developed world often work in low-quality jobs, regardless of education previously attained in their country of origin. It then analyses the long-term value of formal host-country education for refugees using the example of forcibly displaced Bosnians who arrived in Austria during the 1992-1995 Bosnian war. By exploiting a unique institutional setting in Austria, where host-country education was assigned quasi exogenously to Bosnian refugees around specific age thresholds, the study shows that host-country education has large implications for the quality of refugee employment. The 22 years of Austrian microcensus data allow me to also track these over time and in the long-run. The discount on Bosnian education is large in the beginning and declines over time for men but not for women, suggesting that host-country degrees are particularly important to groups that faced cultural barriers to quality employment in their country of origin. I show that even after more than two decades of stay in the host country, Bosnian refugees who hold formal degrees from Austria were still significantly less likely to work below their educational attainment and earned more than comparable Bosnians with similar degrees from Bosnia.

1.2.3 Third Essay

In the third essay, I take a forward looking approach to migration that originates from developing countries by zooming in on the role of global food prices, one of the climate-related factors that have become significantly more volatile over the past decades. I show that poor agricultural households' ability to send one of their members as a labour migrants is a positive function of locally-relevant global crop prices. The reason is that migration is still a costly investment to these households and income-increasing rises in global food prices act as a migration enabler; vice versa, drops in global crop prices decrease the ability of households to send a labour migrant. I show that the order of magnitude of a standardised price effect is approximately one third of the standardised effect of a local weather shock. Unlike positive weather shocks, which mostly facilitate internal rural-urban migration, positive income shocks through rising producer prices only increase migration to neighbouring African countries, likely due to the simultaneous decrease in real income in nearby urban areas. The data further allows me to rule out conflict onset as an alternative mechanism of the positive association between global prices and household out-migration: I show that, while higher producer prices indeed induce conflict, conflict does not play a role for the household decision to send a member as a labour migrant.

1.2.4 Fourth Essay

The fourth essay then takes the perspective of a destination country. It is motivated by previous findings in the literature that show a negative association between migrant density and the willingness to redistribute income among natives. The paper analyses the impact of the large, unexpected and spatially heterogenous migration wave from Central and Eastern Europe to England following the accession of the new-joiners to the EU in 2004 on local authority spending and revenue. I apply a difference-in-differences estimation strategy and find that migrants did not have an effect on local authorities' total service provision per capita. Once I zoom in on the different expenditure items, I find that local authorities that experienced relatively larger migration inflows saw their spending on means-tested social care services decrease in relative terms, while spending on education services increased. To further explain these results, I specifically disentangle a potential change in local preferences for redistribution from more mechanical changes to revenue and spending patterns brought about by migrants' distinct characteristics. Examining changes in local Council compositions and internal migration flows in response to the arrival of outsiders, I find no evidence that the observed spending shifts were driven by a change in the local willingness to redistribute income. Rather, the results suggest that, due to migrants' young age at the time of arrival, migration following the 2004 EU enlargement mechanically decreased social care demand and alleviated some of the pressure social care spending in England faces.

1.3 Reflections and policy implications

In each chapter of this PhD thesis, I explain in detail how the respective work relates to existing research, how my own work attempts to close gaps in the literature and what policy implications can be derived from the results. This section provides a general reflection on the broader lessons I learnt from the four years I spent on researching the economics of migration.

The first one is the realisation that the impacts of migration on destination countries are rarely one-dimensional and often don't actually reflect intrinsic differences of migrants compared to native populations. To illustrate with an example why this is important, consider the study by Bell et al. [2013] that assesses the impact of migration on local area crime rates in the UK. The authors show that refugees, but not other migrants, increase local property crime rates. Without adding nuance to such findings, these results can easily lead the reader to conclude that asylum seekers possess characteristics that make them intrinsically more likely to commit crime. The policy implications could then have severe negative consequences for asylum seekers. The authors acknowledge this and discuss the limited labour market access among asylum seekers in the UK as a potential explanation for their results, which would imply that the negative behavioural response is not necessarily specific to migrants. In a later study, Couttenier et al. [2019] indeed show that access to the labour market can significantly reduce crime rates among asylum seekers. It is therefore important to reflect carefully on institutional arrangements and mechanisms before assigning outcomes to group-level behavioural differences.

As a consequence, much of the time spent on the research within this thesis was dedicated to thinking about mechanisms that could explain the obtained findings. For example, in chapter 5, I show that the large wave of Central and Eastern European migrants to the UK following the 2004 EU enlargement decreased the means-tested social care spending in areas more heavily affected by these inflows. Without further investigation, one could interpret this outcome as reflective of a lower local preference to share with outsiders as found in previous cross-country studies. However, when analysing the local preferences channel, I find no evidence that changes in the local preference for redistribution is actually the driving force behind the reduction in spending on means-tested items. Rather, the findings suggest that the younger age of migrants relative to the native population reduced demand for local social care services. Note that this fundamentally changes the policy implications: While a decrease in the local willingness to redistribute following a large migration wave would raise doubts about the sustainability of welfare

systems, the demographic channel actually implies an alleviation of the pressure these systems face.

A second related point is the careful thought about counterfactuals in migrant integration studies. A range of studies have recently started assessing the impact of various refugee-related polices on employment outcomes. Consider the carefully designed study by Hainmueller et al. [2016] who analyse the effect of waiting times during the asylum procedure in Switzerland (when asylum seekers had very limited access to the labour market) on employment outcomes in the year after asylum had been granted. The authors find that a longer waiting time reduces subsequent employment rates among refugees. While, as a standalone finding, this is certainly a cause for reducing the time asylum procedures take to reach a decision, it makes the -in their setting, likely correct- assumption that asylum seekers stay mostly idle during these procedures. However, a policy counterfactual in which asylum seekers are offered host-country education while waiting for their asylum decisions would likely produce results very similar to those found in the study: A longer time before entering the labour market would then partly reflect the amount of host-country human capital they accumulate, which would in turn likely lead to higher reservation wages and lower employment rates in the early years after arrival. Thus, by slightly changing the institutional setting, the interpretation of results can change dramatically. In fact, one of the key takeaways from the research conducted in chapter 2 and 3 of this thesis is that the notion of employment rates as a success metric for refugee integration should at the very least be taken with a grain of salt when measured shortly after the arrival of refugee migrants. Headline employment figures often mask that early employment is limited to low-income employment which comes at the expense of the accumulation of host-country human capital. In chapter 2, I further show that choosing to work in low-quality professions in the early years after arrival may partly reflect distortions, with unknown long-term consequences for both host-country welfare and for refugees themselves.

The third lesson relates to the importance of quality data. A large amount of time within this PhD project was dedicated to identifying and cleaning suitable data sources. It strikes me as almost ironic that some of the most heatedly debated issues such as the integration of refugees into host-country labour markets or migration related to climatic events in developing countries suffer from a lack of quality data to assess these phenomena. While it is certainly the job of a researcher to work with incomplete data sources, data limitations may very well limit the policy implications that can be derived from analyses. For example, in chapter 4, I construct a panel of backward-reported household

out-migration from a cross-section of household data to study the impact of global prices and local weather on these rates. I am able to show that poor agricultural households' ability so send one of their members as a labour migrant depends on locally-relevant global prices and local growing conditions, likely due to the high upfront costs of migration. However, due to the lack of data on household income over time, it is not possible to study the exact global crop price - household income elasticity. Yet, the elasticity would be helpful to know for development aid agencies that want to assess the scope of agricultural households potentially becoming trapped when global crop prices become more volatile in a future characterised by a changing climate.

Finally, I believe that the economics of migration discipline has a lot to learn from other literatures outside the field of economics. For example, in a recent review of the forced migration literature, Becker and Ferrara [2019] argue that the forced migration experience can affect integration outcomes in the destination country through major disruptive life events displaced persons endure. When addressing this gap in the literature in chapter 2, I learnt that literatures in sociology and psychology had dealt with questions related to refugees' mental health and behavioural consequences in quite some detail and with nuance. Engaging with other fields is difficult, and the detailed exchanges I had with a psychiatrist on how victimisation may affect behaviour were often slow due to differences in jargon and conditioned thinking patterns. However, I also learnt that breaking through these established patterns is particularly rewarding and the ideas that emerge from interdisciplinary dialogue can open up new research agendas.

Chapter 2

Barriers to humanitarian migration, victimisation and integration outcomes: Evidence from Germany

2.1 Introduction

One of the key features of humanitarian migration flows from developing into developed regions of the world is the significant risk these journeys entail for individuals who embark on them. According to the International Organisation for Migration's (IOM) Missing Migrant database, around 15,000 migrants perished in the Mediterranean Sea alone when trying to reach the territory of European Union (EU) member states between 2015 and 2019. Asylum seekers who survive the perilous journey often do not come out unscathed: They are subjected to violent acts on their journey carried out by escape agents and border enforcement agencies, with detrimental consequences for their physical and mental health [Albahari, 2018, Arsenijević et al., 2017, 2018]. Against the backdrop of a subdued economic and societal integration of newly-arriving humanitarian migrants in the EU [Brell et al., 2020], the potentially negative consequences of these victimisation events for future life trajectories of affected individuals -and thus the welfare of hosting countries- has increasingly found its way into the political debate. For example, the European Commission [2020] notes that "mental health is critical to migrants' integration" [p. 3] and "especially refugees, may be at higher risk of developing mental health problems due to [...] difficulties encountered during their migration journey" [p. 13].

In this paper, we explicitly analyse the potentially disruptive role of asylum seeker

victimisation during their journey to safety for their economic activity in Germany, the main destination country of asylum seeker inflows into the EU. To analyse this link, we deploy novel refugee survey data collected in the aftermath of the large inflows of asylum seekers into the country following the 2015 European migration crisis. We use this data to construct a physical and a financial victimisation indicator for each refugee based on detailed questions regarding negative events individuals experienced along the journey to Germany. We use these indicators to first study the effect of victimisation on life satisfaction, health satisfaction, overall economic activity, labour force participation and education in the destination. We then link the survey to administrative employment biography data to further explore how both physical and financial victimisation affect refugees' employment trajectories upon arrival in the host country.

Our results show that physically victimised asylum seekers in particular suffer from significantly lower life satisfaction as well as poorer physical and mental health upon arrival. We find that physically victimised refugees are less likely to invest into host-country education but take up employment faster than non-victimised and financially victimised migrants. This leads to the intuitively paradoxical finding of a higher employment rate among physically victimised refugees vis-à-vis other refugees in the early years upon arrival, an adjusted gap reaching 4.4 percentage points 31 months into refugees' stay in Germany. We show that this additional employment among the physically victimised is driven by marginal and part-time work that is characterised by a lower level of income relative to other refugees. We further document larger financial hardship among the financially victimised but do not find distortionary effects on the human capital investment decision for this subgroup. Our findings thus suggest that, in line with previous research, physical victimisation has stronger effects on affected individuals' behaviour [Dolan et al., 2005, Mahuteau and Zhu, 2016, Johnston et al., 2018].

To conceptualise our findings, we draw on evidence from sociology, psychology and economics literatures. Evidence from sociology and psychology documents a "loss of future directedness" caused by potentially traumatic victimisation events [Beiser, 1987, Hauff and Vaglum, 1993a, Hunkler and Khourshed, 2020, Sagbakken et al., 2020]. As physically victimised refugees adopt a more pessimistic outlook on life and discount their future more heavily, they tend to invest less into host-country specific education and are more likely to take up low-skill employment soon after arrival. Similarly, the experimental economics literature shows that time preferences can be affected by extreme events linked to violence, making victimised individuals more impatient in their decision-making [Voors et al., 2012, Callen et al., 2014, Jakiela and Ozier, 2019, Brown et al., 2019]. While these

effects are not directly distinguishable from a general decline in mental well-being in our setting, one of the key strengths of our study is that the granular survey data allows us to rule out a number of competing theories that could explain our findings. These range from institutional mechanisms in-built into German asylum procedures to mechanisms related to financial hardship and other behavioural changes caused by victimisation events. Since the first survey interviews were conducted close to the arrival date of asylum seekers, our findings do not face the standard issue of reverse causality when linking victimisation and the resulting decline in mental well-being to individuals' labour market attachment [Brown et al., 2010, Kassenboehmer and Haisken-DeNew, 2009]. The relatively large number of observed individuals further allows us to split our sample along a range of dimensions. We show that all our findings are robust to restricting the sample to recognised refugees, hold for the subgroup of Syrians who receive protection with near certainty, and are driven by both males and females.

Our identification strategy relies on the random nature of victimisation among asylum seekers. We identify and group four main empirical challenges to a causal interpretation of the effect of victimisation on labour market and health outcomes. First, a non-random selection at the country of origin at different expected victimisation rates may lead specific subsets of asylum seekers to choose Germany as their destination (selection at origin bias). If these groups show unobserved characteristics that also causally impact integration outcomes, estimated coefficients on the effect of victimisation will be biased. A second source of potential bias may arise from the fact that only those who succeed in making it to Germany are observed in our sample. This group may represent a sub-group of, for instance, particularly motivated or resilient individuals (survivor bias). Third, victimisation itself may be correlated with other unobserved factors that determine the economic integration in Germany (omitted variable bias). Finally, a potential bias may arise if the willingness to report victimisation events is related to economic outcomes. To address selection at origin and survivor bias, we show that our baseline results are robust to a wide range of empirical specifications that limit variation to narrowly defined fixed effects categories related to the country of origin of asylum seekers, interacted with their precise time of departure and arrival. We show that, conditional on the geographical origin and the timing of migration, baseline differences between physically and non-physically-victimised groups indeed disappear, leaving us with a sample balanced along a wide range of observable characteristics. By combining these balance tests with coefficient stability tests developed by Oster [2019], we are able to credibly rule out omitted variable bias as a driver of our results in all linear regressions. The rich survey data further allows us to add

precision to our estimates by selecting controls for a wide range of pre-migration and carefully selected post-migration characteristics based on both economic theory and machine learning techniques. Taken together, throughout our analysis, we thus compare victimised and non-victimised refugees with similar pre-migration and selected post-migration characteristics, who originate from the same country, migrated in the same year-month, took the same migration route and are part of the same arrival cohort. To rule out misreporting as a source of bias in our analyses, we conduct several systematic tests on the sample of refugees who agreed to answer journey-related questions, showing that neither the willingness to answer nor social desirability are likely to taint the obtained results.

Our study adds to the existing literature in a number of ways. We primarily add to the literature that links refugee victimisation to their economic behavioural response in the host country [Couttenier et al., 2019, Hunkler and Khourshed, 2020, Hauff and Vaglum, 1993a]. Unlike previous literature, our data allows us to explicitly focus on events asylum seekers endured during their journey as opposed to their country of origin, an important distinction for the design of asylum policies. Since victimisation of asylum seekers is interconnected with the choice of external border policies [Arsenijević et al., 2017, 2018], we further contribute to the growing literature on how asylum-seeker specific policies shape refugee labour market integration [Damm, 2009, Battisti et al., 2016, Hainmueller et al., 2016, Marbach et al., 2018, Zwysen, 2019]. One of the main takeaways of our study is that a rapid labour market integration as a general success metric for integration outcomes should be treated with care; higher victimisation rates may contribute to a relatively swift uptake of employment, nevertheless distorting the timing of labour market entry. We further add to the recently developing stream of literature that links crime victimisation to labour market outcomes more generally [Bindler and Ketel, 2019]. We show that this link is context-specific and the perilous journey asylum seekers go through may not be easily compared to other settings. Finally, we also add to the much broader literature on violence and the human-capital investment decision by providing further evidence that experiencing traumatic events in conflict and high-crime settings lowers the willingness to invest in education [Blattman and Annan, 2010, Shemyakina, 2011, Leon, 2012, Akbulut-Yuksel, 2014, Koppensteiner, 2021].

The remainder of this paper proceeds as follows. Section 2.2 discusses the conceptual framework that links the victimisation experience to economic activity outcomes in the destination country in more detail. Section 2.3 discusses our data sources. Section 2.4 introduces the estimation strategy and our approach to dealing with self-selection and survivor bias in detail. Section 2.5 shows the main results and section 2.6 discusses their

robustness. In section 2.7, we test alternative hypotheses that could explain our findings. Section 2.8 provides a concluding discussion.

2.2 Related literature and research hypotheses

Some evidence exists on the direct disruptive effect of victimisation events on affected individuals' future economic outcomes for the general (Western) population without an explicit focus on migrants or refugees. These studies unequivocally find negative consequences of victimisation regarding labour force participation, employment, earned income and welfare dependency [Bindler and Ketel, 2019, Ornstein, 2017, Velamuri and Stillman, 2008]. Other studies, primarily on developing countries, have documented distortions to the human-capital investment decision following potentially traumatic events in conflict-related or high-crime settings [Blattman and Annan, 2010, Shemyakina, 2011, Leon, 2012, Akbulut-Yuksel, 2014, Koppensteiner, 2021]. A larger literature is concerned with the effect of victimisation on health and mental well-being outcomes, one of the structural mechanisms linking victimisation events to disruptions in economic outcomes. For example, Mahuteau and Zhu [2016] find that physical victimisation decreases subjective well-being. Dolan et al. [2005] estimate a measure of 'loss in quality-adjusted life years' to quantify the cost to victims and show that rape, followed by other serious physical assault decreases victims' quality-adjusted life years the most. Johnston et al. [2018] focus on the effect of crime victimisation on life satisfaction as an encompassing measure of mental well-being and find that physical assault and events that lead to a major worsening of affected individuals' financial situation have the strongest negative effect on life satisfaction for both men and women.

Despite these clear findings for the general population, it is crucial to note that the situation of asylum seekers who arrive in their host-country is unique and the extent to which general studies apply to migrants' victimisation experience requires a careful reflection. Unlike samples of victimised individuals drawn from the general population or other migrant groups, forcibly displaced migrants all start their economic activity trajectory at zero upon arrival. This has major implications for their choice set with respect to the host-country labour market and potential distortions to these choices caused by a victimisation event. The majority of refugees originate from countries where educational attainment is not regarded as equivalent to education in economically advanced countries and refugees often lack proof of their formal degrees. Therefore, refugees from less developed countries

face the decision to i) either join the labour force immediately, accept a discount on their educational attainment and take up low-skill, low-income employment or ii) invest into host-country specific human capital to have access to better paid employment later on [Duleep and Regets, 1999, Cortes, 2004].

Evidence from both sociology and psychology literatures suggests that victimisation experiences have lasting effects on future-oriented planning. In a study closest to ours, Hauff and Vaglum [1993a] analyse a cohort of 145 Vietnamese refugees in Norway one year and three years after arrival. The authors find that those who experienced more potentially traumatic events were more likely to be in the labour force but tend to invest less into host-country specific education. In a separate study on the same cohort, the authors find a persistently worse mental health status among those who experienced traumatic events even seven years after arrival vis-à-vis the non-traumatised [Hauff and Vaglum, 1993b]. In combination, these findings suggest that experiencing trauma under extreme conditions may either trigger a behavioural change on the short-term versus long-term trade-off independent of the effect of victimisation on mental health or as a direct consequence of it. In a more recent study, Hunkler and Khourshed [2020] evaluate the integration outcomes of 252 Syrian refugees in Germany's federal state of Bavaria and broadly confirm the findings by Hauff and Vaglum [1993a]. The authors find that refugees who experienced potentially traumatic events in their country of origin or on their journey to Germany show better German language skills but their labour market outcomes do not differ from the non-victimised, a curious finding considering the clearly established link between language skills and economic integration success [Dustmann and Soest, 2001].

The tendency of refugees to adopt a more present-oriented mindset following extreme potentially traumatic episodes can be traced back in the psychology literature until Beiser [1987]. The author finds that refugees who went through extreme events while fleeing their country have a shortened sense of future, a potential explanation for favouring early employment over long-term educational investment. In a more recent study, Sagbakken et al. [2020] interview 78 asylum seekers in Norway and find that a "loss of future directedness" leads traumatized asylum seekers to "withdraw into passivity" [p. 1]. Research on self-harm and suicide among refugees show that the lack of future orientation also manifests itself into higher rates of suicide in these segments of the population, especially when having endured extreme traumatic events [Goosen et al., 2011, Hawton, 2009, Lerner et al., 2016].

These findings that relate traumatic events refugees endure to a lack of future-oriented

planning find further support in recent experimental studies in the economics literature. While classic economic models assume that preferences are stable over time and unaffected by life-time experiences [Stigler and Becker, 1977], recent experimental studies suggest that individuals' risk aversion and time preferences can indeed be affected by extreme events linked to violence [Voors et al., 2012, Callen et al., 2014, Jakiela and Ozier, 2019, Brown et al., 2019], natural disasters [Eckel et al., 2009, Page et al., 2014, Callen, 2015, Cameron and Shah, 2015, Cassar et al., 2017, Hanaoka et al., 2018, Beine et al., 2020], health shocks [Decker and Schmitz, 2016], and financial and macroeconomic shocks [Guiso et al., 2018, Jetter et al., 2020, Malmendier and Nagel, 2011, Kettlewell, 2019]. The time preferences have in turn been found to affect decisions regarding borrowing [Meier and Sprenger, 2010], savings [Thaler and Benartzi, 2004] and financial literacy [Meier and Sprenger, 2013] among adults, and investment and human capital acquisition among the younger [Sutter et al., 2013, Cadena and Keys, 2015, Kemptner and Tolan, 2018]. Our analysis relates to this literature by indirectly measuring the time preferences of victimised versus non-victimised individuals.[1] This possible interpretation is based on the assumption that individuals indirectly reveal their time preferences by engaging in certain activities [DellaVigna and Paserman, 2005]. Individuals who attach higher value to long-term rewards are more likely to pursue activities that entail an immediate cost (such as investing in human capital) but which have delayed payoffs (access to higher quality employment in the future). On the other hand, impatient individuals are more likely to engage in activities which have immediate benefits (such as low-income employment) and delayed costs (lack of access to higher quality employment in the future).

In summary, we derive three hypotheses regarding the effect of victimisation events asylum seekers experience on their journey to Europe:

Hypothesis 1: Asylum seekers' victimisation during the journey to Europe leads to lower mental well-being and general health at destination.

We start by testing hypothesis 1 since it allows us to (i) validate the severity of victimisation events on affected individuals' well-being and (ii) motivate hypothesis 2a. For our hypotheses related to the effect of victimisation on economic outcomes in the destination we then rely on a revealed preferences argument. We test two competing hypotheses:

[1] Voors et al. [2012], for instance, find that individuals who were exposed to greater levels of violence display more altruistic behaviour, are more risk-seeking, and have higher time discount rates. Similarly, Cassar et al. [2017] conduct a series of experiments in rural Thailand and find that the 2004 tsunami led to long-lasting increases in risk aversion, pro-social behaviour, and impatience

Hypothesis 2a: Due to their relatively lower mental health, victimised refugees are expected to be less attached to the labour market and invest less into education than non-victimised refugees.

Hypothesis 2a follows directly from the literature that studies the effect of victimisation on mental health and the literature linking lower mental health to adverse economic and education outcomes in the general population. However, the unique situation of refugees at the time of arrival in their destination leads to a second competing hypothesis regarding their economic integration:

Hypothesis 2b: Due to a "loss of future directedness", victimisation during the journey leads refugees to favour joining the labour force and to take up low-income employment rather than investing into host-country human capital compared to the benchmark of non-victimised refugees.

Since both mechanisms put forward in hypotheses 2a and 2b may be at play simultaneously, we posit that the direction of the estimated effect of victimisation on short-term labour market outcomes will capture whichever channel is stronger. Our approach to testing the proposed mechanisms has two limitations. The first limitation pertains to the measurement of future directedness, which we do not observe directly in the data. Any results in line with the mechanism underlying hypothesis 2b thus requires us to test all alternative hypotheses that conceivably link asylum seeker victimisation during the journey to early integration outcomes in the same way. These hypotheses relate to the institutional setting in Germany, potential differences in financial hardship caused by a different intensity of engagement with people smugglers during the journey and potential differences between the victimised and the non-victimised in the intention to stay in Germany in the long term. We discuss these in detail in section 2.7. The second limitation pertains to the detail with which our approach unveils the underlying causal chain. Our data does not allow us to unambiguously infer if the loss of future directedness is caused by the decrease in mental health or if the potentially traumatic experience directly changes time preferences.

2.3 Data and definitions

This section contains a description of all data sources and introduces the main variables of interest and the outcomes. In the final subsection 2.3.4, we then turn to a first descriptive analysis of differences between victimised and non-victimised asylum seekers to both fill

a gap in the literature and as a first motivation of our empirical strategy.

2.3.1 IAB-BAMF-SOEP refugee survey

The main data source for all our analyses is the IAB-BAMF-SOEP refugee survey. It is an extension to the established German Socio-Economic Panel (GSOEP) survey through an ad hoc module for the target population of asylum seekers and refugees. The sampling frame of the IAB-BAMF-SOEP survey is the German Central Register of Foreign Nationals. The survey has a panel structure with a total of 12,311 interviews carried out in three waves in 2016, 2017 and 2018 on 6,763 individuals.

The survey contains a wide range of baseline information on pre-migration characteristics and detailed information on individual and household characteristics, including on respondents' health. It also provides information on the time of displacement in the country of origin, within-country information on the province of origin and the time of arrival in Germany. Most importantly for our purposes, respondents are asked detailed questions on experiences they went through during the journey from their country of origin to Germany. 3,742 individuals, 55.2% of the total sample, agreed to provide information on these experiences. Our effective working sample for which we have all relevant information, including all necessary control variables and additional outcomes, consists of 2,314 individuals aged between 18 and 65.

Within these questions, our main interest lies on the survey question 'During the journey or escape, did you experience one or more of the following?' which allow surveyees to choose one or more answers from a list of negative experiences. Based on their responses, we create a **binary physical victimisation indicator** taking the value 1 if an individual was subjected to sexual abuse, physical attacks, incarceration or shipwreck (or any combination of these). We further create a **binary financial victimisation indicator** taking the value 1 if an individual was subjected to financial fraud, extortion, robbery or blackmail, or any combination of these. The summary statistics for the two victimisation indicators are shown in table 2.1.

Variable	Mean	Std. Dev.	Min.	Max.
Experienced robbery	0.133	0.34	0	1
Experienced extortion	0.155	0.362	0	1
Experienced fraud	0.287	0.452	0	1
Financial victimisation	0.39	0.488	0	1
Experienced sexual harassment	0.017	0.129	0	1
Experienced shipwreck	0.137	0.344	0	1
Experienced physical attack	0.134	0.341	0	1
Experienced incarceration	0.201	0.401	0	1
Physical victimisation	0.359	0.48	0	1
N		2314		

Table 2.1: Physical and financial victimisation indicator

We note that by design of these indicators, individuals may experience both a financial and a physical trauma. Reassuringly, the correlation between these two (r=0.3) is sufficiently low not to be a cause for concern in our regression analyses. We further note that some migrants experienced more than one victimisation event, but nevertheless model our preferred indicators as binary for two main reasons. First, the majority of migrants experienced one victimisation event. Only 11% of all individuals in our sample experienced more than one physical victimisation event and 15% experienced more than one financial victimisation event. Second, there is no clear guidance in the victimisation literature on the correct functional form of the relation between our outcomes of interest and multiple victimisation events individuals experience during their flight to safety that lasted for 40 days on average. We explore this question further by relaxing the assumption of no additional behavioural effect when individuals experience additional victimisation events in section 2.6.1.

Important to our analysis, from the second wave of the survey onwards, individuals who were willing to answer questions regarding their escape journey were explicitly asked which route they took to reach their destination. We assign their answers to the five main migration routes: (1) The Eastern Mediterranean sea route, (2) the Central Mediterranean route, (3) the Western Mediterranean route, (4) the Eastern Mediterranean land route, (5) the Eastern Land border route and (6) travelling directly to Germany by plane. Since the survey questions on the route taken contain many missing values in the first wave of the survey, we impute the routes using an additional source of information. Individuals were invited to report on a virtual map all locations they passed through on their migratory journey from their country of origin to Germany. We use this data and apply the method developed by Guichard et al. [2021] to extract the geo-referenced points, infer the migration route and classify these to match the five first routes.[2] Applying these

[2]The authors start by assigning the geo-coded points to all countries, and define a sequence of coun-

methods allows us to recover route information for 77 % of the sample. We assign the remaining 23 % to a seventh category (7), no route information available.

We then draw on the psychiatric and health economics literature reviewed in section 2.2 and use four main indicators to measure mental well-being and the general health of victimised and non-victimised refugees. Following Johnston et al. [2018], we use life satisfaction measured on a scale from 1 to 10 (with 10 being the highest), an encompassing measure of mental well-being, as our primary outcome of the mental health effect of victimisation. We complement this measure with a measure of self-assessed health (1-10), a mental component score (MCS) and a physical component score (PCS). The MCS is a standardised mental health index, based on six questions related to emotional and psychological problems and how these impact on daily activities [Nübling et al., 2006]. The factor loadings of the questions are used as weights and the score is then normed to range from 0 to 100. Similarly, the physical health index is constructed based on six questions related to physical health. We provide all details on the construction of the MCS and the PCS in section 2.C of the appendix. The four measures are summarised in table 2.19 of appendix 2.A for the first observation available of each individual in the panel, together with the backward-reported measures of pre-migration life-satisfaction and pre-migration self-reported health.

Finally, for the economic integration outcomes, our main interest lies on three measures: First, the general willingness to engage in economic activity, defined by those in the labour force and those in education or training. In the second step, we zoom into the generic economic activity indicator and break it up into those in the labour force and those in education and training to uncover potential distortions to the timing of entering the labour force caused by victimisation. We complement our main analyses with an analyses on employment rates, further split up into full-time, part-time and marginal employment, and net monthly income. These are shown in table 2.20 of appendix 2.A for the last observation available of each individual in the panel.

tries for each migration route. Secondly, they identify the last country before an individual entered the Schengen area and the first location in the Schengen zone. A path is assigned to (1) the Eastern Mediterranean sea route if the last non-Schengen country was Turkey and the first Schengen country was Greece; (2) the Central Mediterranean route if the last non-Schengen country was Egypt, Libya, Tunisia, or Turkey and the first Schengen country was Italy or Malta; (3) the Western Mediterranean route if the last non-Schengen country was Morocco or Algeria and the first Schengen country was Spain or France; (4) the Eastern Mediterranean land route if the last non-Schengen country was Turkey and the first EU country was Bulgaria; (5) the Eastern Land border route if the last non-Schengen country was Romania, Ukraine, or Belarus, and the first Schengen country was Poland, Slovakia, or Hungary.

2.3.2 IAB integrated employment biographies

We use the IAB integrated employment biographies (IEB) to complement the employment labour market survey questions with more reliable individual administrative records. The IEB consist of all individuals in Germany who are characterised by at least one of the following employment status: employment subject to social security (in the data since 1975), marginal part-time employment (in the data since 1999), benefit receipt according to the German Social Code III or II, officially registered as job-seeking at the German Federal Employment Agency or (planned) participation in programs of active labour market policies (in the data since 2000). The social security notifications are filled by the employer for each employment relationship. Unique establishment identifiers allow collapsing the employee level (e.g. Establishment History Panel (BHP)) and to link workers, establishments and surveys (e.g. Linked-Employer-Employee-Data (LIAB)). The IEB is a comprehensive data set with no attrition and daily precision.

While the IEB data can only be linked to the survey questions for a 70% subset of our sample for a total of 1625 individuals and we therefore rely on employment outcomes in the survey data in our primary analyses, the more precise job market data with a longitudinal character adds three major components. First, it provides us with the exact date of the first job refugees took up in Germany, allowing us to address the question of the timing of employment uptake in greater detail. Second, the linkage allows us to follow refugees even when they leave the survey, mitigating attrition concerns. Finally, it also allows us to obtain information on refugees' pre-survey (un)employment history.

2.3.3 Further data sources

All survey information are further linked to the Uppsala Conflict Data Program and to the Syrian Shuhada Martyr Revolution database on the province-month of displacement level. These databases report aggregate number of fatalities by province and month between 2011 and 2019 for all countries of origin found among the refugee population in Germany. We use these data to calculate control variables for potentially traumatic experiences in the origin, which we proxy in two steps. First, we construct a province-specific conflict-related death count in the province of origin, defined as the twelve-months rolling average of conflict related fatalities prior to departure. As argued by Aksoy and Poutvaara [2019], simple province-level death counts may not adequately capture conflict intensity, as all variation in the variable may come from few historically war-ridden countries with

substantially different institutional settings. We follow their approach and calculate a measure of conflict intensity (CI) the following way:

$$CI_{c,t-\mu} = \frac{\sum_{m=1}^{12} TotalDeaths_{c,t-\mu-m}}{12}.$$

For each country and month $t-\mu$, we then calculate the median conflict intensity, M, of all provinces and create three categories: "No conflict", for all individuals departing at $t-\mu$ for whom $CI_{c,t-\mu} = 0$; "Low conflict", for all individuals departing at $t-\mu$ for whom $CI_{c,t-\mu} < M$ and "High conflict", for all individuals departing at $t-\mu$ for whom $CI_{c,t-\mu} \geq M$.

Thus, the conflict intensity measure is calculated based on within-country conflict variation over time.[3] The calculated variables are summarised in table 2.21 of appendix 2.A.

2.3.4 Victimisation along the migration route to Germany

To the best of our knowledge, no systematically gathered statistics exist on the hardship asylum seekers endure during their journey to Europe. In 2014, the International Organisation of Migration's (IOM) "Missing Migrant" project started tracking fatalities along the main migration routes but the database does not feature information on survivors. Two findings from observational studies are relevant to our study. First, victimisation is frequent and leaves asylum seekers mentally scarred. Arsenijević et al. [2017] evaluate data on 992 asylum seekers on their way to Europe collected from Médecins sans Frontières' mobile mental health clinics in Serbia in 2015 and 2016. Almost a third of the asylum seekers transiting through Serbia was classified as physically traumatised by the medical staff. Second, perpetrators of violence are both state actors and people smugglers asylum seekers rely on to cross borders and the Mediterranean Sea [Albahari, 2018, Arsenijević et al., 2017]. Less is known about individual and circumstantial factors that lead to victimisation on the journey to Europe and to what extent these events are of random nature. Anecdotal evidence strongly suggests that crime victimisation is not limited to groups traditionally seen as vulnerable. For example, Arsenijević et al. [2018] show that young male migrants are often the victims of physical assault and suffer severe mental health problems due to these events.

[3] An alternative solution to the problem Aksoy and Poutvaara outline would be to use the death count measure in combination with country by year-of-departure fixed effects.

To shed more light on victimisation events endured by asylum seekers during their journey to Europe, we start out by presenting more detailed descriptive statistics comparing group characteristics of non-victimised, physically victimised and financially victimised migrants in appendix 2.B. In summary, table 2.22 of appendix 2.B suggests that the group composition of non-victimised and victimised migrants differ along a range of dimensions. The victimised groups show some signs of being less likely to show features generically considered vulnerable. Victimised migrants are on average younger, less likely to be female and have fewer children. The victimised groups also show characteristics that can be classified as greater determination to reach Germany: On average, both physically and financially victimised asylum seekers spend larger amounts on their escape, take more perilous sea routes to Germany instead of arriving by plane, travel for longer and are more likely to travel alone. The differences we observe around the geography of victimisation are intuitive: For example, asylum seekers from Syria are likely to be at a higher risk of victimisation on their journey to Germany than an asylum seeker from Kosovo simply due to the fact that Syrians have to travel a longer distance and spend more time exposed to potential perpetrators of violence. Similarly, victimisation rates will differ by travel route and means of transport: For instance, flying into Germany from Syria is safer than crossing the Eastern Mediterranean Sea in a boat provided by an escape agent. Both the geography and route chosen may therefore reflect underlying financial constraints, which in turn correlate with other socio-economic characteristics. Finally, the timing of migration may further contribute to the observed differences as it likely reflects selection effects at different expected victimisation rates: For instance, relatively vulnerable migrants may choose not to migrate when the expected risk of getting victimised is high.

To systematically analyse these issues, we proceed as follows. We start by discussing the empirical challenges that are behind these group level differences in section 2.4. Based on the empirical strategy we develop, we then show in 2.5 that when balancing on the geographic origin, the migration route taken and the timing of migration, observed average group characteristics between victimised and non-victimised are much more similar.

2.4 Empirical strategy

The key assumption of the analyses in this study is that the victimisation of individuals during their journey to Europe, when conditioned on the right set of covariates, is a random event that affected individuals have no control over. Four empirical problems

pose a risk when trying to interpret the unconditional effect of individual-level victimisation on integration outcomes: First, a non-random selection at the country of origin at different levels of expected victimisation may lead specific groups of asylum seekers to choose Germany as their destination (selection at origin effect). If these groups show characteristics such as unobserved motivation that also causally impact on integration outcomes, estimated coefficients on the effect of victimisation pick up bias. Second, a potential bias may arise from the fact that only those who succeed in making it to Germany are observed in our sample. This group may represent a sub-group of, for instance, particularly motivated or resilient individuals (survivor bias). Third, victimisation itself could be correlated with other unobserved factors that determine the integration in Germany (omitted variable bias). If, for instance, unobserved intelligence lowers the probability of victimisation and simultaneously improves economic decision-making in the destination, coefficients estimated on the victimisation indicator could pick this up. Finally, there is a potential bias arising from misreporting of victimisation events.

The following subsections explain our strategy to deal with the different sources of potential bias in detail. We start by laying out our baseline specification that estimates the effect of victimisation of asylum seekers along the route to Germany on their economic integration in section 2.4.1. The subsequent subsections 2.4.2, 2.4.3 and 2.4.4 explain the extensions. Subsection 2.4.5 then addresses the threat of a potential systematic bias arising if the reporting of victimisation events was indeed linked to the outcomes.

2.4.1 Baseline specification

The rich set of background information available from the IAB-BAMF-SOEP survey data allows us to control for pre-migration and selected post-migration characteristics that are usually unobserved in migration studies. This allows us to reduce some concerns related to potential unobserved variable bias.

To identify the effect of victimisation events occurring during the journey to the destination on economic integration outcomes, we start out by estimating the following em-

pirical model:

$$
\begin{aligned}
Y_{i,c,t,\mu,a,f} =& \gamma_1 PhysicalVictim_i + \gamma_2 FinancialVictim_i \\
&+ \zeta BaselineCharacteristics_{i,t} + \eta PreMigCharacteristics_{i,\mu} \\
&+ \theta_1 ConflictIntensity_{i,\mu} + \theta_2 Route_{i,\mu} + \theta_3 MSM_{i,t} + \theta_4 MSMsq_{i,t} \\
&+ \theta_5 AsylumStatus_{i,t} + \alpha_c + \beta_a + \delta_f + \tau \hat{\lambda}_i + \epsilon_{i,c,t,\mu,a,f}
\end{aligned} \tag{2.1}
$$

where $Y_{i,c,t,\mu,a,f}$ captures the health or economic integration outcome of interest for individual i from country of origin c, who was interviewed at time t, had left their origin at time μ, arrived in cohort a and resides in the German federal state f. Both γ_1 and γ_2 are the coefficients estimated on the variables of interest, $PhysicalVictim_i$ and $FinancialVictim_i$.

$BaselineCharacteristics_{i,t}$ is a vector of individual level characteristics. It includes refugees' age, age squared, a dummy for female refugees, a measure of mental resilience and willingness to take risk and a set of dummies referring to individuals' educational attainment (ISCED-2011 categories).[4] We include mental resilience as a control since more resilient individuals might be better able to cope with distressful life events. Willingness to take risks is controlled for to account not only for self-selection into migration but also because of its confounding effect on time preferences. For the labour force participation, education and health related outcomes we include two categorical variables related to the residence of the spouse and to the location of the children.[5] For the health related outcomes, we further include a continuous variable reflecting the satisfaction with living arrangements, measured on a scale from 1 to 10. We note that it is unclear to what extent the willingness to take risk, mental resilience and satisfaction with living arrangements could themselves be affected by victimisation events. To account for the possibility of these variables constituting bad controls, we show in appendix 2.F that our main results do not depend on their inclusion.

[4]The mental resilience scale is based on the procedure suggested by Jacobsen et al. [2017]. The scale is based on the responses to four questions: "I try to think of how I can change difficult situations.";"No matter what happens to me, I think I have my reactions under control."; "I think I can develop further if I deal with difficult situations."; "I actively seek ways to balance out the losses that have affected my life." The response scale ranges from 1 (disagree) to 7 (fully agree) and the resilience variable is the average of these responses. The willingness to take risk variable is based on the question "How do you rate yourself personally? In general, are you someone who is ready to take risks or do you try to avoid risks?" The response scale ranges from 0 (not prepared to take risks at all) to 10 (Prepared to take risk)

[5]The residence of the spouse variable contains the following categories: Single; the spouse lives in the same household, the spouse lives in a different household in Germany; the spouse resides abroad. The variable related to the location of children contains the following categories: No children; all children live in the same household; some children live in a different household

$PreMigrationCharacteristics_{i,\mu}$ is a vector of individual level pre-migration characteristics measured at departure μ. These include information on the economic situation, knowledge of German, employment experience and backward reported measures of health and life satisfaction before migration. It further includes a dummy variable taking the value one for individuals who arrived alone in Germany, and zero otherwise. Finally, it includes a dummy variable that takes the value 1 for individuals who had help from relatives or acquaintances when moving to Germany.

$ConflictIntensity_{i,\mu}$ measures the conflict intensity in the province of origin measured on within-country conflict variation as explained in section 2.3.3. We include the conflict intensity control in the baseline specification for two reasons: First, to account for previous findings from psychological research showing that the individual-specific response of victimisation depends on previous traumatic experiences [Yehuda, 2002, Breslau et al., 2008]. Second, to account for selection into migration dynamics at different levels of push factors at the origin [Aksoy and Poutvaara, 2019, Guichard, 2020].

The categorical variable $Route_{i,\mu}$ indicates the migratory route taken as defined in section 2.3.1. In theory, conditioning victimisation estimates on the migratory route could constitute a bad control if all routes that can possibly be taken were understood as a choice set. This is unlikely to be the case for the reason that routes are only distinguished on a high level in our data and are to the largest extent determined by the geography of the country of origin, and partly by the time of forced displacement. Conditioning on the route taken is nevertheless informative beyond geography for two reasons. First, it allows to explicitly account for travel by plane where migrants face a significantly lower victimisation risk. Travel by plane is further indicative of high wealth and social status in the country of origin. Second, conditioning on the migratory route allows to distinguish sea from land travel. Travelling by land makes victimisation events such as shipwreck unlikely to occur, while the risk of others such as experiencing physical assault by border enforcement agencies may increase when crossing borders by land instead of by sea. For the above reasons, we include the migratory route taken as a control variable. We nevertheless show in appendix 2.F that all our main results do not depend on this choice.

MSM is the number of months a refugee spent in Germany and $MSMsq$ is its squared term.

$AsylumStatus_{i,t}$ is a time varying individual level characteristic, measured at the time of the survey t. It is a fixed effects term with four categories: "Asylum granted", "Tem-

porary suspension of deportation", "Request to leave Germany" and "Decision pending". Only the first two give refugees unrestricted access to the labour market in Germany, an institutional feature we discuss in more detail in section 2.7.1.

$\hat{\lambda}_{i,c,t}$ is an estimated Heckman correction term if the outcome of interest is the (log of) income, which requires individuals to be employed. To identify the first stage of the Heckman correction, the excluded variables are the residence of the spouse and the location of the children.

α_c captures country of origin fixed effects and β_a refers to cohort of arrival fixed effects, which we measure in three categories, 2013-2014, 2015 and 2016-2017, in the baseline specification. δ_f reflects the German federal state in which the refugee is residing at the time of the survey. We include the latter purely to gain precision as refugees are not allowed to move freely and are assigned to a federal state initially upon arrival [Aksoy et al., 2020].

Finally, $\epsilon_{i,c,t,\mu,a,f}$ is an error term. All standard errors are obtained using delete-cluster jackknife methods due to the estimated (unknown) $\hat{\lambda}_i$ if the Heckman correction term is included.[6] In the results section we refer to the set of controls included in equation 2.1 as "Baseline controls".

Since the survey has a longitudinal dimension but our variable of interest is non-time-varying, we estimate equation 2.1 in two ways. First, we estimate the model as a cross-section, (i) only using the first observation available of each individual to study the effect of physical and financial victimisation on outcomes related to the (mental) health and well-being. We use the first observation available to study these outcomes for two reasons. First, when refugees were interviewed for the first time, they had spent only 19 months in Germany on average. Thus, their mental well-being related outcomes can be expected to still be affected by negative experiences during the journey to Germany or in their home country. Second, the potential reverse causality problem of mental well-being and employment is minimised [Brown et al., 2010, Kassenboehmer and Haisken-DeNew, 2009]: Only 9.2% of our sample were employed in the month prior to the interview when first surveyed. We then (ii) use the last observation in the sample to study the effect of victimisation on economic integration outcomes. At this point, individuals had spent an average of 31 months in the country and 21.8% were employed. While the

[6]The industry standard here is to cluster-bootstrap standard errors. Due to the large amount of fixed effects in the regression, bootstrapping becomes computationally impossible. Note that the delete-cluster jackknife method produces slightly more conservative standard errors than bootstrapping [Efron, 1992]. The estimated Heckman correction term is only included if statistically significant in the second stage.

average difference between the first and last observation of each individual is therefore only 12 months -and thus concerns about potential sample attrition due to selective return migration are minimised- the additional variation we gain in our outcomes of interest is considerable.

In a second step, we then also exploit the panel variation in the data and estimate a (individual i) random effects model under the assumption that $corr(\epsilon_{i,f,c,t}, X) = 0$.[7] The large number of time constant variables in the model - including the set of fixed effects related to the time of migration and the origin of individuals - makes this key assumption of a random effects model plausible in our setting [Wooldridge, 2010]. We note that, since all asylum seekers naturally start their stay in Germany as economically inactive and the likelihood of engaging in economic activity then increases over time, the panel estimates that capture average effects over time are not directly comparable to the cross-sectional estimates based on only the final observation of each individual.

2.4.2 Self-selection into migration to Germany at the origin

One of the key concerns in equation 2.1 as noted in section 2.3.4 is the bias the coefficients γ_1 and γ_2 may pick up due to potential selection effects at the country of origin. In our setting, selection at the origin relates to the concept of migration at different levels of expected victimisation risk. Since (past) cohort level victimisation rates can be understood as a indirect measure of the (expected) journey risk, these findings require careful consideration in our empirical strategy. We first note that limited evidence on the self-selection of forced migrants at the origin at different expected journey risk levels has started to emerge in recent academic literature. Aksoy and Poutvaara [2019] provide suggestive evidence that intended destinations change when country-specific risk levels are altered through stricter migration policies, with potential consequences for the cohort composition. The authors further show that a higher conflict intensity at the origin leads to self-selected migration of more highly educated asylum seekers, in particular among female migrants. It follows that at an increased expected journey risk, which can also be understood as an increase in migration cost, self-selection may become even more salient.

Despite controlling for individuals' willingness to take risk, the route travelled and conflict intensity at origin around the time of migration in our baseline specification, the concerns around self-selection into migration at the origin can most efficiently be tackled

[7] Thus, this specification assumes that the individual-specific residual is uncorrelated with the explanatory variables X.

by the use of a large set of fixed effects related to the country of origin and granular time of migration. The detailed information on migrants' journeys we obtain from our surveys allows us to add interactive country-of-origin by year-month-of-migration fixed effects, $\kappa_{c,m}$, to equation 2.1. We label this specification the *Fixed Effects* specification in our regression tables and, due to its efficiency in eliminating self-selection dynamics, refer to it as our preferred specification throughout this study.

2.4.3 Survivor bias

Not all forcibly displaced migrants who decide to embark on the journey to Germany make it to their preferred destination. If the selection of asylum seekers who eventually reach their targeted destination is a random subset of those individuals that initially decided to migrate there, selection during the journey would not be an empirical concern when studying the effect of victimisation during the journey on integration outcomes as long as self-selection at the origin is accounted for. If granular country of origin by time of departure fixed effects are integrated as suggested in section 2.4.2, these would absorb random shocks to migration cohorts: Even if specific cohorts then show relatively higher or lower victimisation rates, the restriction to within-departure-cohort variation would prevent estimates on variables of interest to pick up systematic bias.[8]

However, changes in the difficulty of the journey may potentially have non-random effects on the arrival cohort composition, even when narrowly conditioning on the selection at the origin at different points in time. We refer to this empirical issue as survivor bias. In theory, such change in the composition of asylum seeker arrival cohorts can influence not only the probability of victimisation but also their performance on the German labour market. While we are able to mitigate this concern by controlling for a large set of observable characteristics, unobservable compositional arrival cohort changes may introduce bias in regressions of integration outcomes on individual level victimisation experiences.

To the best of our knowledge, no previous research exists that could inform our empirical strategy with regards to survivor bias and the extent to which it is a concern in our setting. Empirically, we therefore address the issue of survivor bias by deploying a large set of dyadic departure - arrival fixed effects in robustness tests. In addition to our preferred specification suggested in section 2.4.2, we thus estimate a model that includes

[8]The effect would be visible in changes in the number of arrivals at any given time but since our focus is on the comparisons at the individual level, this does not impact our estimation strategy.

year-month of arrival fixed effect, interacted with the year-month of departure and region of origin, $\varrho_{c,m,a}$, thereby limiting the variation to even narrower categories within the region of origin, the month of departure and the month of arrival. Due to the inevitable loss in degrees of freedom, we include broader categories of migrants' origin as shown in table 2.22 in these interactive fixed effects, in addition to the country of origin fixed effects.

2.4.4 Further methods to address omitted variable bias

The rich set of background information available from the IAB-BAMF-SOEP survey data allows us to control for pre-migration and selected post-migration characteristics, mitigating potential unobserved variable bias. Deploying a large set of fixed effects as explained in 2.4.2 and 2.4.3 allows us to further capture institutional changes and cohort characteristics that could otherwise bias the coefficients estimated on our variables of interest. A standard drawback of including a large set of fixed effects is a loss of statistical precision, which could lead to discarding estimated effects that show no significance at conventional statistical levels. One way to more efficiently deal with the bias-variance trade-off while mitigating concerns around theory-driven model specifications is the use of machine learning methods for data-driven model selection. To test the sensitivity of our results to our modelling choice, we therefore follow Belloni et al. [2014] who develop a 'post-double selection' (PDS) method to estimate separate least absolute shrinkage and selection operator (LASSO) regressions to find predictors of the selection equation(s) and the outcome equation when many potential controls are available. Their procedure thus allows us to consider all interactions and non-linearities that are not considered in theory-driven specifications. Hence, instead of simply including the control variables as in 2.1, the PDS approach considers all possible interactions and non-linearities between the control variables and the country of origin and time of departure fixed effects. Furthermore, we consider a set of additional control variables we did not include previously due to their large number of missing values. These variables include: The use of a smuggler, the log of the cost of the smuggler, the financing of the journey (e.g., sale of assets, borrowing, savings, among others), means of transportation used to reach Germany (e.g., boat, car, foot, train, plain), self-reported reason for migrating (e.g., persecution, discrimination, economic, among others), having stayed in another country for three or more months before coming to Germany, and the log of the duration of the journey in days.

We apply the PDS methodology to our two variables of interest, $PhysicalVictim$ and $FinancialVictim$ and proceed in a post-triple selection procedure: In a first step, we

estimate the final outcome equation using LASSO, excluding both the $PhysicalVictim$ and the $FinancialVictim$ variable to obtain a first set of LASSO selected controls. In a second step, we estimate the probability of physical and financial victimisation separately on the same set of explanatory variables to obtain a second and third set of LASSO selected controls. Finally, in a third step, we estimate a linear model similar to our baseline equation 2.1 that includes the union of control variables selected by LASSO in the various steps. A more technical explanation of the procedure is provided in appendix 2.D.

Despite controlling for a wide range of relevant pre-migration and selected post-migration variables as well as accounting for selection effects at various stages of the migratory journey, omitted variable bias cannot be ruled out with absolute certainty. We therefore follow Oster [2019] to test if controlling for observables, and the stability of the estimated coefficients on our variables of interest when conditioning on these, mitigates bias arising from unobservable factors. The technique suggested by Oster [2019] thus informs us about the salience of omitted variable bias in our setting: It provides an estimate of the relative importance of unobserved factors compared to those we do observe. We implement her methodology for our preferred specification and first define a value for R_{max}, the hypothetical R-squared value of a fully specified model which includes all relevant control variables. Oster [2019] recommends a value of $R_{max} = 1.3\hat{R}$, where $\hat{R}$ is the R-squared value obtained from the estimated model; we choose a slightly more conservative approach and set the R_{max} to $1.5\hat{R}$. In a second step, we then obtain the δ that informs us about the relative importance of omitted variables compared to those variables we condition our estimates on: For example, a value of $\delta = 1$ means that unobserved factors would have to be as important as those that are observed for γ_1 and γ_2 of equation 2.1 to equal zero instead of the obtained estimate. We report the estimated δ for our preferred specification in section 2.5.3.

2.4.5 Reliability of self-reported victimisation

Finally, one of the concerns when using sensitive survey data on victimisation is the reliability of responses. In our setting, two particular sources of bias need to be ruled out, which our data allows us to do under weak assumptions.

The first concern relates to a potential link between the employment status of respondents and the willingness to answer questions on victimisation events - which can be either

positive or negative. First, it is conceivable that employed individuals feel more comfortable to share their victimisation experience if employment is indicative of a relatively more stable life that allows putting distance between the present and past experiences. We refer to this concern as "willingness to answer bias". Second, unemployed individuals could be more willing to answer the questions on victimisation events to justify the difficulties they experience to integrate into the labour market. We refer to this concern as "social desirability bias" [Krumpal, 2013]. The social desirability would bias estimates in the opposite direction.

The willingness to answer bias and social desirability bias are largely addressed by the panel structure of our data. Questions on victimisation events are only asked in the first interview, when a large share of refugees had only arrived recently. At the time of the first interview, the average time since migration was one year and eight months. Only 9.2% of refugees in our sample were employed at that time, a number we verify using IAB administrative data for the subsample of refugees for which this information is available. Thus, employment status is unlikely to make a large difference in the willingness to answer the victimisation question. To further mitigate concerns around these sources of bias, we regress the willingness to answer the victimisation questions on the employment status reported during the first interview, using our preferred specification as outlined in section 2.4.2.[9] The results of this exercise is shown in the left panel of table 2.25. None of the estimated coefficients is significantly different from zero at any conventional level, leading us to accept the null hypotheses of no "willingness to answer" and no "social desirability" bias.

The second potential source of bias relates to a potential systematic misreporting of victimisation events: We rely on the truthful reporting of victimisation experiences among those who agreed to answer questions related to these negative experiences. Without further investigation, it is conceivable that individuals who reported victimisation experiences are the ones who feel comfortable to report these because the experienced events were not traumatic to them. On the other hand, those who did not report victimisation events could have been most severely affected by traumatic events. In this case, the victimisation indicators would paradoxically capture the opposite of what they intend to measure.

We first note that the structure of the survey questions largely reduces this concern. Before any journey-related or victimisation-related questions are asked, individuals are

[9]We use our full sample of respondents who agreed and did not agree to answer the victimisation-related questions.

explicitly confronted with the following introductory question: "Next, we have a few questions about the experiences connected with your escape. Some of the questions will be about negative experiences. Would you like to answer questions about this subject or would you prefer not to answer these questions?" Thus, individuals are given the opportunity not to answer questions related to their own victimisation experiences. Once individuals agree to reply to this set of questions, the assumption that they will reply to these questions truthfully becomes plausible.

Based on the answers respondents give regarding their willingness to reply to questions about negative experiences during the journey, we can further test the representativeness of the subsample of those who were willing to reply to these questions. The right panel of table 2.25 shows the results of regressing individuals' employment status in the last available survey wave on the willingness to respond to journey-related questions, again using our preferred specification. We do not find the willingness to respond to victimisation-related questions to be a significant predictor of employment outcomes. Thus, we conclude that the underlying characteristics that predict whether or not an individual is willing to respond to sensitive survey questions in our setting are either unrelated to our outcomes of interest or fully accounted for by the included covariates.

2.5 Results

In this section, we first discuss the main results which link victimisation on asylum seekers' journey to Germany to the (mental) well-being outcomes and the economic integration in subsection 2.5.1. We then turn to employment outcomes in subsection 2.5.2 and present our test results for the significance of unobserved confounders in subsection 2.5.3.

Two tradeoffs need to be balanced in our regression analyses. First, the large amount of fixed effects we introduce based on our rationals presented in sections 2.4.2 and 2.4.3 absorb a lot of the variation in our outcomes. They effectively deal with bias due to selection effects but lead to a loss in statistical precision. Second, some of the information on the arrival time we require to construct the dyadic time of departure - month of arrival categorical variable is not available for all individuals, decreasing the sample size. Our analyses throughout the paper suggest that these tradeoffs are best balanced by adding fixed effects on the country of origin, the time of departure and their interaction as suggested in section 2.4.2. We therefore refer to this specification as our preferred specification. All other results derived from specifications discussed in the previous section

are nevertheless shown for completeness and robustness. Simplified regressions that do not include any control variables, as well as alternative specifications excluding covariates that could potentially be considered outcomes of victimisation events are presented in appendix 2.F for all main results. We further show the estimated coefficients for all control variables included in our preferred specification in appendix 2.H.

Before turning to the main results, we return to the question of group-level differences between victimised and non-victimised migrants raised in section 2.3.4. Table 2.2 shows a conditional balance test, under the null hypotheses that individual-level characteristics do not predict victimisation events once the selection dynamics outlined in section 2.4 are accounted for. To test these multiple hypotheses, the physical victimisation indicator and the financial victimisation indicator are regressed on a set of backward reported pre-migration indicators respectively, conditional on their geographical origin, the time of migration (and their interaction term), and the migration route as suggested by our preferred specification. Physical victimisation is further conditioned on having experienced financial victimisation, while financial victimisation is conditioned on the experience of physical victimisation. The regression outcomes are shown in column (1) and (2). Column (3) and (4) then additionally condition on the the time of arrival.

	Physical Victim. (1)	Financial Victim. (2)	Physical Victim. (3)	Financial Victim. (4)
Female	0.00228	-0.0182	0.00346	-0.0172
	(0.0245)	(0.0268)	(0.0246)	(0.0268)
Age	0.00183	-0.00335	0.00158	-0.00385
	(0.00630)	(0.00662)	(0.00631)	(0.00663)
Age squared	-6.21e-05	1.31e-05	-5.93e-05	1.77e-05
	(8.21e-05)	(8.71e-05)	(8.22e-05)	(8.72e-05)
Willingness to take risk	0.00110	0.00325	0.00110	0.00338
	(0.00310)	(0.00327)	(0.00310)	(0.00327)
Resilience	-0.00163	-0.000257	-0.00165	-0.000376
	(0.00310)	(0.00329)	(0.00310)	(0.00330)
Life satisfaction BFM	-0.00158	0.00842*	-0.00167	0.00803*
	(0.00473)	(0.00484)	(0.00473)	(0.00486)
Health satisfaction BFM	-0.00340	-0.0187***	-0.00330	-0.0185***
	(0.00514)	(0.00527)	(0.00515)	(0.00527)
Employed BFM	0.0265	0.0583**	0.0271	0.0579**
	(0.0270)	(0.0285)	(0.0270)	(0.0285)
Education: Secondary	0.0238	0.0914***	0.0229	0.0882***
	(0.0280)	(0.0293)	(0.0280)	(0.0293)
Education: Vocational	-0.0379	-0.0199	-0.0392	-0.0218
	(0.0408)	(0.0427)	(0.0408)	(0.0427)
Education: Tertiary	-0.0471*	0.0636**	-0.0490*	0.0581**
	(0.0263)	(0.0291)	(0.0265)	(0.0291)
German skills BFM: Good	-0.165*	-0.193**	-0.159	-0.171*
	(0.0984)	(0.0879)	(0.0991)	(0.0887)
Economic Situation BFM < Avg	0.0231	-0.0103	0.0225	-0.00998
	(0.0289)	(0.0303)	(0.0289)	(0.0303)
Friends helped to move	-0.0404	-0.0829***	-0.0399	-0.0868***
	(0.0253)	(0.0268)	(0.0253)	(0.0267)
Arrived alone	0.00759	-0.00336	0.00817	-0.00503
	(0.0266)	(0.0279)	(0.0266)	(0.0279)
Observations	2,314	2,314	2,314	2,314
R-squared	0.229	0.158	0.230	0.163
C.origin*Departure FE	Yes	Yes	Yes	Yes
Migration Route FE	Yes	Yes	Yes	Yes
Cohort FE	No	No	Yes	Yes
Physical Victimisation	Yes	No	Yes	No
Financial Victimisation	No	Yes	No	Yes

Huber-White SE; *p<.1; **p<.05; ***p<.01

Note: BFM stands for backward reported "before migration" information. The term FE indicates fixed effects. The term cohort refers to the year-month of arrival in Germany. Willingness to take risk, Life satisfaction BFM and Health satisfaction BFM are measured on a scale from 1 (low) to 10 (very high). The baseline category for the education variable is below secondary education. The baseline category for the dummy variable "German skills BFM: Good" is no or very limited German skills before migration. The baseline for the dummy variable "Economic Situation BFM < Avg" is "Economic Situation BFM > Avg" and refers to the economic situation of individuals compared to the population in their country of origin. "Friends helped to move" and "Arrived alone" are dummy variables capturing if migrants received help from friends during their migratory journey and whether they arrived alone respectively.

Table 2.2: Conditional balance test

The results show a balanced sample for the physically victimised (columns 1 and 3). We note that, on average, tertiary educated individuals and those who spoke German before their forced displacement are slightly less likely to experience physical victimisation. This difference is only statistically significant at the 10 percent level (1), is no longer visible for the German language indicator when conditioned on cohort fixed effects (3) and could be the result of multiple hypotheses testing. Nevertheless, our estimations will routinely control for these variables. Experiencing financial victimisation, on the other hand, is unsurprisingly correlated with indicators that identify the wealth status of individuals (columns (2) and (4)). Higher levels of education and employment before migration both predict financial victimisation events, likely due to the necessary condition of initial

financial endowment to suffer financial losses. Our data allows us to control for a wide range of pre-migration wealth indicators, mitigating the issue. Similarly, individuals who could rely on friends to help with their migratory journey - and are therefore less likely to be involved with escape agents - show a lower probability of victimisation on average. Finally, life and health satisfaction are further significant predictors of financial victimisation, albeit in opposing directions. Our suggestive interpretation is that life satisfaction is correlated with financial victimisation via its link to financial endowment, whereas individuals of relatively worse pre-migration health are easier targets for criminal actors. In sum, we conclude that while the sample of physically victimised migrants is balanced along a wide range of individual level pre-migration characteristics once conditioned on geography and the timing of migration, financial victimisation events occur less randomly. While our data allows us to control for a large set of potentially confounding variables to mitigate the problem, we deploy several additional tests, outlined in 2.4.4, to study the significance of unobserved factors that could bias our results.

2.5.1 Main results

We first turn to the results on the encompassing measures of mental well-being in table 2.3.

	Cross Section				Panel Data	
	Benchmark	Fixed effects	Dyadic FE	PDS	Benchmark	Fixed effects
Panel A: Life Satisfaction	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	-0.150	-0.170*	-0.236*	-0.214**	-0.106	-0.161**
	(0.0935)	(0.0978)	(0.129)	(0.1000)	(0.0703)	(0.0741)
Financial victim.	-0.0868	-0.105	-0.178	-0.133	-0.0696	-0.0996
	(0.0914)	(0.0978)	(0.126)	(0.0974)	(0.0687)	(0.0734)
R-squared	0.289	0.351	0.438			
Panel B: Health Satisfaction	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	-0.214**	-0.281**	-0.230	-0.263**	-0.176**	-0.261***
	(0.109)	(0.117)	(0.148)	(0.115)	(0.0848)	(0.0909)
Financial victim.	-0.222**	-0.181	-0.299**	-0.280**	-0.151*	-0.153*
	(0.106)	(0.114)	(0.148)	(0.113)	(0.0816)	(0.0876)
R-squared	0.269	0.331	0.414			
Observations	2,261	2,261	1,711	2,109	4,864	4,864
Baseline Controls	Yes	Yes	Yes	Some	Yes	Yes
C.origin*Departure FE	No	Yes	Yes	Some	Yes	Yes
R.origin*Departure*Arrival FE	No	No	Yes	Yes	No	No

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable captures self-reported life satisfaction or self-reported health satisfaction on a scale from 1 to 10, 10 being the highest. Columns (1) to (4) use observations corresponding to the date of the first interview conducted, 19 months after arrival on average. Panel data results are derived from a random effects specification. The term FE indicates fixed effects. PDS refers to the post double-selection LASSO regressions. The term departure refers to the year-month of forceful displacement from the home country. Arrival FE refers to the year-month of arrival in Germany. C.origin is the country of origin. R.origin is the wider region of origin by continent. Since the precise information on the year-month of arrival is not available for all individuals, the dyadic FE regressions are estimated only on the subset of individuals where this information is available. The PDS is estimated on the same sample as the Fixed Effects regression, but drops singleton observations.

Table 2.3: Health and Life Satisfaction After Migration

Throughout specifications, the effect of physical victimisation on life satisfaction at the time of arrival in Germany is negative (panel A). All estimated coefficients are statistically significant at the 10% level or close to that threshold. In our preferred specification (column (2)), the magnitude of the coefficient is 0.170, which corresponds to a decrease of approximately 10% in the standard deviation of the measure. The estimated coefficients are stable across specifications. Financial victimisation also shows a negative effect on life satisfaction across specifications, but the estimated coefficients are smaller and less precisely estimated.

Turning to the self-reported health outcomes in table 2.3 confirms these results (panel B). In our preferred specification of column (2), physical victimisation decreases self-reported health by 0.281 points ($p<.05$), again corresponding to approximately 10% in the standard deviation of the measure. The estimated coefficient is again stable across specifications. The negative effect of financial victimisation on self-assessed health is also visible across specifications, albeit being of smaller magnitude in out preferred specification (column 2). In summary, two main findings confirm the effect of victimisation on mental well-being and health established in previous studies on the general (non-refugee) population [Mahuteau and Zhu, 2016, Dolan et al., 2005, Johnston et al., 2018]. First, the event of victimisation has a negative effect on both measures. Second, physical victimisation leaves a stronger effect on the overall well-being of individuals and a marginally stronger effect on individuals' self-assessed health compared to financial victimisation. In appendix 2.C, we split the health measure into a physical and mental component and show that the overall result is driven by a combination of both.

Table 2.4, panel A, then turns to the effect of physical and financial victimisation on general economic activity, defined as those in the labour force or pursuing host-country specific education.

	Cross Section				Panel Data	
	Benchmark	Fixed effects	Dyadic FE	PDS	Benchmark	Fixed effects
Panel A: LFP or Education	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	0.0378**	0.0384**	0.0362	0.0278	0.0314**	0.0384***
	(0.0177)	(0.0189)	(0.0253)	(0.0191)	(0.0135)	(0.0148)
Financial victim.	-0.0106	-0.0076	0.0128	-0.0098	0.0084	-0.0018
	(0.0174)	(0.0183)	(0.0244)	(0.0186)	(0.0134)	(0.0145)
R-squared	0.245	0.311	0.386			
Panel B: LFP	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	0.0546***	0.0579***	0.0525**	0.0442**	0.0402***	0.0513***
	(0.0179)	(0.0192)	(0.0258)	(0.0192)	(0.0137)	(0.0150)
Financial victim.	-0.0087	-0.0089	0.0161	-0.0056	0.0120	-0.0035
	(0.0178)	(0.0188)	(0.0250)	(0.0187)	(0.0136)	(0.0147)
R-squared	0.251	0.315	0.392			
Panel C: Education and Training	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	-0.0266**	-0.0296**	-0.0293*	-0.0249*	-0.00783	-0.0128
	(0.0116)	(0.0127)	(0.0168)	(0.0129)	(0.0080)	(0.0096)
Financial victim.	-0.0002	-0.0007	0.0022	0.0051	-0.0067	-0.0033
	(0.0118)	(0.0125)	(0.0156)	(0.0127)	(0.0079)	(0.0091)
R-squared	0.115	0.210	0.301			
Observations	2,314	2,314	1,754	2,159	4,864	4,864
Baseline Controls	Yes	Yes	Yes	Some	Yes	Yes
C.origin*Departure FE	No	Yes	Yes	Some	Yes	Yes
R.origin*Departure*Arrival FE	No	No	Yes	Yes	No	No

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals in the labour force or in education (Panel A), in the labour force (Panel B) or pursuing host-country education or training (Panel C). LFP stands for Labour force participation. Columns (1) to (4) use observations corresponding to the last interview conducted, 31 months after arrival on average. Panel data results are derived from a random effects specification. The term FE indicates fixed effects. PDS refers to the post double-selection LASSO regressions. The term departure refers to the year-month of forceful displacement from the home country. Arrival FE refers to the year-month of arrival in Germany. C.origin is the country of origin. R.origin is the wider region of origin by continent. Since the precise information on the year-month of arrival is not available for all individuals, the dyadic FE regressions are estimated only on the subset of individuals where this information is available. The PDS is estimated on the same sample as the Fixed Effects regression, but drops singleton observations.

Table 2.4: Economic Activity

We do not find a negative effect of victimisation during the journey on general economic activity in the cross-sectional regressions (1) to (4) estimated on individuals who, on average, had spent 31 months in Germany. In our preferred specification of column (2), physical victimisation has a positive effect on being economically active ($p<.05$). These results are further confirmed in the panel regressions (5) and (6). Since these panel regressions include observations of the same individuals at an earlier point in time (when individuals had spent 19 months in Germany on average) and the cross-sectional regressions do not include these, the stability of coefficients when comparing the cross-sectional and panel results is noteworthy; the findings suggest that the gap in economic activity between physically and non-physically victimised already opens up at least 19 months after arrival. No such effect can be found for the financially victimised.

To shed more light on the drivers of this finding, table 2.4, panel B, shows the results of the regressions of labour force participation on our victimisation measures. The coefficients on labour force participation show a strong, precisely estimated positive association of physical victimisation and joining the labour force across all specifications. In our preferred specification of column (2), this effect is estimated at 5.8 percentage points ($p<.01$). The effect remains visible and of only slightly smaller magnitude in the panel specifications (5) and (6), suggesting that physically victimised individuals indeed join the labour force sooner upon arrival. We do not find the same association between financial victimisation and labour force participation, where the estimated effect is close to zero across all specifications.

Table 2.4, Panel C, then shows the regression outcomes of the effect of victimisation on pursuing host-country education and training. By design, the results complement those of Panel A and Panel B. Physical victimisation significantly decreases the propensity to pursue host-country specific education or training across all specifications. In our preferred specification of column (2), this negative effect reaches 3.0 percentage points, a very sizeable decrease considering that the total share of refugees in our sample pursuing education or training stands at 7.8 percentage points 31 months after arrival (see table 2.20). Three points are further noteworthy: First, the panel regressions of column (5) and (6) suggest that this effect only becomes visible after some time in the country. The estimated coefficients are smaller in magnitude and no longer distinguishable from zero when including observations on the same individuals closer to arrival. Second, the lower share of physically victimised refugees in education and training does not entirely close the gap to the higher labour force participation of the same group shown in table 2.4. In combination, these two observations suggest that the barriers to pursuing host-country

education are higher and education opportunities open to refugees require more time to search for than joining the labour force. Finally, the coefficients estimated on the financial victimisation indicator again show no effects across all specifications.

In summary, our findings indicate that the physical victimisation event i) increases the propensity to join the labour force early on and ii) decreases the propensity to pursue host-country education and training. The results in the three panels of table 2.4 thus suggest that the act of physical victimisation leads to a distortion in the timing of labour force entry, which appears to dominate the more general well-being related effects that would likely lead to lower economic activity rates. We interpret these findings as supportive of the "loss of future directedness" hypothesis.

In section 2.5.2, we further put this hypothesis to the test by explicitly i) considering whether the higher labour force participation rates indeed result in higher employment rates, ii) analysing the type of employment victimised individuals engage in vis-à-vis the non-victimised and iii) shedding more light on the timing of first employment when comparing the different groups.

2.5.2 Employment outcomes

Our findings shown in section 2.5.1 strongly support the idea that the act of physical victimisation reduces future-oriented thinking among the affected. This section more explicitly considers the consequences of an early labour force entry: If the "loss of future directedness" was indeed a relatively stronger driving force within the group of the physically victimised, we would expect a lower reservation wage and thus a relatively higher take-up of readily available low-income employment among those who experienced physical victimisation events on their journey to Germany.

Table 2.5 shows regression results with different types of employment rates as the dependent variable for the full sample of refugees. We focus on the whole population as the underlying population - rather than only those in the labour force - to make these regressions directly comparable to those in table 2.4. We nevertheless report the same employment rates for only those in the labour force in appendix 2.G. We further only report results for our preferred specification in both the cross-section and the panel in all subsequent analyses.

Employment	Any employment		Full-time		Part-time or marginal		Log of income	
	Cross S.	Panel D.	Cross S.	Panel D.	Cross S.	Panel D.	Cross S.	Panel D.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Physical victim.	0.0435**	0.0326**	0.0132	0.00931	0.0303**	0.0236**	-0.163	-0.128
	(0.0190)	(0.0134)	(0.0148)	(0.0104)	(0.0147)	(0.0106)	(0.123)	(0.108)
Financial victim.	0.0027	0.002	0.0059	0.0039	-0.0032	-0.0028	-0.0343	-0.00967
	(0.0185)	(0.0128)	(0.0145)	(0.0098)	(0.0142)	(0.0100)	(0.122)	(0.107)
Observations	2,314	4,864	2,314	4,864	2,314	4,864	408	750
R-squared	0.258		0.239		0.152		0.372	
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for employed individuals, with regressions showing employment outcomes for any employment ((1) and (2)), full-time employment ((3) and (4)) and part-time or marginal employment ((5) and (6)). When employment is split up, the other types of employment are set to zero. All samples use observations corresponding to the last interview conducted, 31 months after arrival on average; Columns (7) and (8) are estimated on the sample of individuals employed after 31 months. Columns (1), (3), (5) and (7) show the results of a cross-sectional regression on the last interview conducted. Columns (2), (4), (6) and (8) are derived from a random effects panel specification. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C.origin is the country of origin.

Table 2.5: Employment

The results in table 2.5 indeed suggest that the higher labour force participation among physically victimised is driven by an increased uptake in part-time and marginal employment vis-à-vis the non-victimised. Column (1) reports the estimated difference in employment rates between physically victimised refugees and the non-victimised for the last observation available of each individual, at an average duration of stay in Germany of 31 months. It is 4.4 percentage points higher than the employment rate among the non-victimised at that point. The panel regression of column (2) - which again includes earlier observations of the same individuals when employment rates were closer to zero and estimates the average employment gap over time - confirms this result. In line with the "loss of future directedness" hypothesis, early employment uptake is characterised by a poor quality of jobs available to refugees. More than two thirds of the gap in the employment rate is explained by employment in part-time and marginal jobs (columns (5) and (6)). Less than one third of the effect is explained by full-time employment, a magnitude that is no longer statistically distinguishable from zero at conventional statistical levels (columns (3) and (4)). Column (7), estimated only on the sample of employed refugees, provides suggestive evidence that 31 months after arrival, these differences already result in a 16% wage gap between the non-physically-victimised and the physically victimised. We note that this difference is likely to increase in the future when the non-physically-victimised complete their training and education, an idea further supported by the slightly smaller coefficient estimated in the panel regression of column (8) where earlier observations are included in the sample.

We have so far not conclusively addressed the question of the timing of employment

uptake of victimised individuals compared to the non-victimised: While comparing the cross-sectional results in column (1), which only contain the last observation of each individual, to the panel results (column (2)) of table 2.5 strongly suggests that joining the labour force early allows for faster access to part-time and marginal employment only, the larger sample size in the panel regressions could also simply add statistical precision to the estimates. To shed more light on the timing of first employment in Germany, we turn to the linked employment biography data, which contains information on the date of first employment. Figure 2.1 shows the unconditional Kaplan-Meier curve of time to first employment, where failure is defined as obtaining employment and the x-axis shows the number of months since arrival in Germany. The analysis is based on a subsample of 1,625 survey respondents who gave their consent to be linked to administrative employment records. Of these individuals, 751 obtained employment at some point over the observation period; we note that this share is larger than the 21.8% in our cross-sectional regressions. The difference is explained by the IEB data extending beyond the last available survey wave. The cross-sectional regressions we presented so far thus correspond to the 31-months-point on the x-axis.

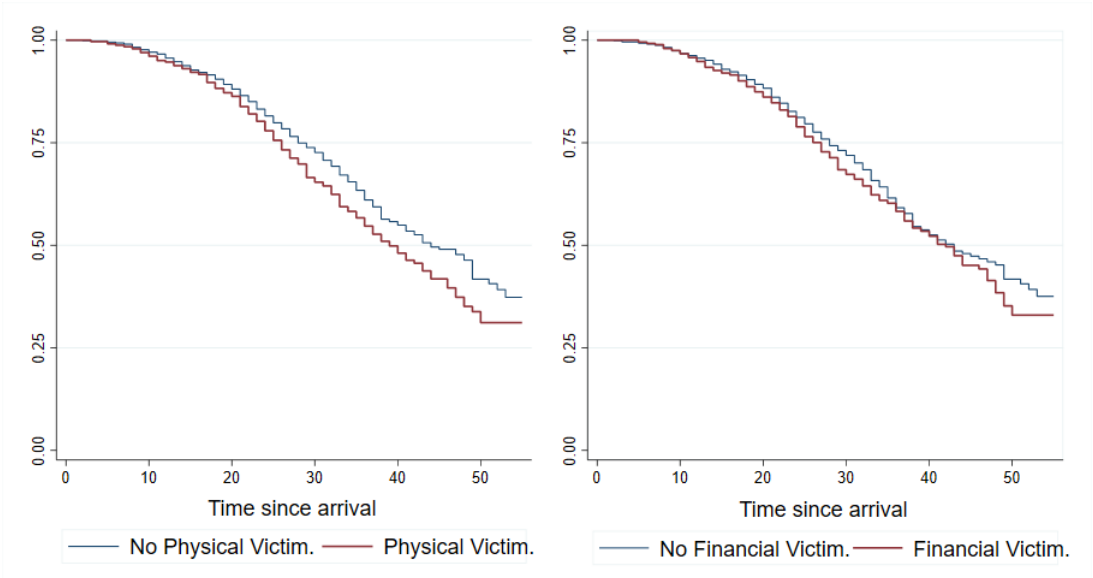


Figure 2.1: Kaplan-Meier estimates of time to first employment

The left hand-side graph in figure 2.1 shows that, compared to the non-victimised refugees, physically victimised refugees obtain employment faster. The gap starts to open around 18 months after arrival, a dynamic we explicitly analyse further in section 2.7.1. The right hand-side graph shows the same comparison for the financially victimised, where we again do not detect any effect. Table 2.6 further reports the estimated output of the simple cox proportional hazard model.

	Time to employment
Physical victimisation	0.209***
	(0.0782)
Financial victimisation	0.00746
	(0.0772)
Observations	1,625

*p<.1; **p<.05; ***p<.01

Table 2.6: Cox proportional hazard model

The parameter estimates show the increase in the expected log of the relative hazard for the physical victimisation and financial victimisation groups vis-à-vis the non-victimised. Exponentiating the parameter estimates shows that the expected hazard, equal to finding employment, is 1.23 times higher for the physically victimised than the non-victimised on average. Although not causal, these results lend further support to the interpretation that physical victimisation events lead to a more present-oriented mindset that attaches more value to immediate payoffs.

We demonstrated in table 2.5 that early employment uptake is generally characterised by part-time and marginal employment (columns 5 and 6). The IEB contains further information on the level of task requirement for a subset of 569 of the 751 employed individuals in our sample, shown in table 2.7.

	No F.V.		F.V.		No P.V.		P.V.		Total	
	%	Obs	%	Obs	%	Obs	%	Obs	%	Obs
1 Unskilled/semi-skilled task	34.4	116	33.6	78	30.9	111	39.5	83	34.1	194
2 Skilled task	46.9	158	45.3	105	47.4	170	44.3	93	46.2	263
3 Complex task	4.7	16	3.9	9	5.6	20	2.4	5	4.4	25
4 Highly complex task	13.9	47	17.2	40	16.2	58	13.8	29	15.3	87

Note: F.V. refers to financial victimisation and P.V. to physical victimisation. Classification of task level according to IEB categories.

Table 2.7: Level of job requirement

The tabulation shows that physically victimised individuals seem to take up jobs with unskilled or semi-skilled task requirements at a higher frequency. Row 1 of table 2.7 shows that the excess share of physically victimised vis-à-vis the non-physically-victimised stands at 8.6 percentage points; on the other hand, the employed share of physically victimised in the skilled task, complex and highly complex task categories is relatively smaller. We note two limitations of this analyses. First, the sample size in most categories is small and should be interpreted with care. Second, since employment is measured at an early stage after arrival, it is likely that the returns to host-country education are not yet fully captured and will pay off at a later stage. Nevertheless, the analysis of skill-requirements

for the jobs performed lends further support to the idea that faster employment uptake among the physically victimised is characterised by low-skill employment.

2.5.3 Testing for the significance of unobserved confounding variables

The key identifying assumption we make in this study is that once we condition on the geography and the timing of migration, and thus self-selection into migration at different expected victimisation levels, the coefficients we obtain on the victimisation variables themselves are unbiased. To strengthen that case, our main regressions further control for a wide range of individual level (pre-migration) socio-economic characteristics to mitigate the risk of omitted variable bias. A further test to assess the likelihood of unobserved confounders as the driving force behind our results was developed by Oster [2019]. In this section, we follow her methodology as laid out in 2.4.4 and estimate how large the effect of unobserved variables would have to be to obtain zero-coefficients on the victimisation indicators in our preferred specification.

Table 2.8 shows the estimated δ values corresponding to the results of our preferred specification in tables 2.3 (Panel A and B column (2)), 2.4 (Panel A, B and C, column (2)), and 2.5 (column (1), (3) and (5)).

Main Outcomes	Life Sat.	Sat. Health	LFP& Educ.	LFP	Educ.
	(1)	(2)	(3)	(4)	(5)
δ (Physical victim.)	1.919	-12.32	1.990	2.697	36.51
δ (Financial victim.)	0.761	2.589	-0.338	-0.315	1.111
Observations	2,266	2,269	2,314	2,314	2,314
R-squared	0.352	0.331	0.306	0.310	0.209
Employment Outcomes	All employ.	FT employ.	Part-Time or marginal		
	(6)	(7)	(8)		
δ (Physical victim.)	1.572	0.588	7.009		
δ (Financial victim.)	0.0389	0.254	-0.338		
Observations	2,314	2,314	2,314		
R-squared	0.297	0.239	0.178		

Notes: The table shows the estimated δ values based on a test for the salience of unobserved confounders following Oster [2019], which show the relative importance of omitted variables compared to those variables we condition our estimates on. The R_{max}, the hypothetical R-squared value of a fully specified model, is set to $1.5\hat{R}$ where $\hat{R}$ is the R-squared value obtained from the respective estimated model. These $\hat{R}$ are obtained from our preferred specification in tables 2.3 (Panel A and B column (2)), 2.4 (Panel A, B and C, column (2)), and 2.5 (column (1), (3) and (5)). LFP means labour force participation. Educ. means education.

Table 2.8: Oster Test

We first turn to the test results obtained on physical victimisation events. All obtained δ values clearly indicate that the explanatory power of omitted variables would have to be very large compared to those variables included in the model for the estimated coefficients on physical victimisation to be zero. For example, in table 2.4, C, column (2), we estimate that refugees who were physically victimised on their journey to Germany were 3 percentage points less likely to be in education or training 31 months after arrival compared to the non-physically victimised. For the obtained coefficients to be zero instead, unobserved variables would have to be 36.5 times larger than those control variables included in the model. The only value below the $\delta = 1$ threshold recommended by Oster [2019] is the obtained coefficient obtained on full-time employment in table 2.5, column (3). However, the coefficient is not statistically significant from zero at any conventional level in our estimation and the result is therefore unsurprising. In sum, all test results suggest that the estimated effect of physical victimisation on integration outcomes are highly robust to omitted variable bias.

Unsurprisingly, almost all obtained δ values on the estimated financial victimisation coefficients lie below the $\delta = 1$ threshold: None of them are significantly different from zero at any conventional level in our main regressions, suggesting no effect of financial victimisation events during the flight to Germany on individuals' well-being and economic integration outcomes.

2.6 Robustness

In this section, we show the results using alternative ways of aggregating the victimisation experiences in subsection 2.6.1 and for specific types of victimisation events in subsection 2.6.2. We then present a range of split-sample regressions to show that the main results are not driven by any particular subsample in the data. Subsection 2.6.3 shows the main results split by major countries of origin, subsection 2.6.4 splits the sample by gender, and subsection 2.6.5 zooms in on different arrival cohorts.

2.6.1 Effects by sum of physical victimisation and sum of financial victimisation events

We start by showing that our results are robust to different specifications of the physical and financial victimisation. In particular, one of our modelling choices in the analyses so

far has been to code the victimisation events as binary indicators. For our integration outcomes, this choice implicitly assumes that once individuals had to endure a physical or financial victimisation event, additional victimisation events do not alter their well-being and behaviour further. In this subsection, we relax this assumption and explicitly consider the precise number of victimisation events individuals endured. Table 2.9 summarises the number of physical (financial) victimisation events by the share of individuals who endured them. The acronym P.V.E. denotes physical victimisation events(s) and F.V.E. financial victimisation events(s).

Physical victimisation			Financial victimisation		
Variable	**Mean**	**Std. Dev.**	**Variable**	**Mean**	**Std. Dev.**
None	0.641	0.48	None	0.61	0.488
1 P.V.E	0.249	0.433	1 F.V.E	0.24	0.427
2 P.V.E	0.091	0.288	2 F.V.E	0.114	0.318
3 P.V.E	0.018	0.134	3 F.V.E	0.035	0.185
4 P.V.E	0.001	0.029			
N	2314		N	2314	

Table 2.9: Summary statistics - number of physical and financial victimisation events

In table 2.10 we first turn to the regression results of our preferred specification using a linear and a squared measure of the number of physical victimisation events, that ranges from zero to a maximum of four, and of the number of financial victimisation events, that ranges from zero to a maximum of three.

	LFP & Educ.	LFP	Educ.	All employ.	FT employ.	Part-Time or marginal
	(1)	(2)	(3)	(4)	(5)	(6)
Number of physical victim.	0.0221	0.0511*	-0.0427**	0.0598**	0.0067	0.0532**
	(0.0280)	(0.0287)	(0.0192)	(0.0293)	(0.0232)	(0.0230)
Number of physical victim. squared	0.0028	-0.0071	0.0133	-0.0195	0.0026	-0.0221**
	(0.0112)	(0.0117)	(0.00830)	(0.0128)	(0.0105)	(0.0101)
Number financial of victim	-0.0004	-0.0084	0.0065	0.0265	0.0256	0.001
	(0.0282)	(0.0288)	(0.0197)	(0.0291)	(0.0230)	(0.0230)
Number financial of victim. squared	-0.0054	-0.0006	-0.0056	-0.0122	-0.0113	-0.0009
	(0.0116)	(0.0118)	(0.0082)	(0.0121)	(0.0097)	(0.0097)
Observations	2,314	2,314	2,314	2,314	2,314	2,314
R-squared	0.306	0.310	0.210	0.258	0.240	0.153
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals in employment/labour force participation/education, and zero otherwise. FT means full-time. LFP means labour force participation. Educ. means education. Results are only shown for our preferred specification, corresponding to column (2) in the main result tables. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C.origin is the country of origin.

Table 2.10: Number of victimisation experiences (continuous)

The results confirm our main results but their interpretation changes. Column (2) now shows that any *additional* physical victimisation event increases the probability of joining the labour force by 5.1 percentage points, an effect again driven by the take

up of marginal and part-time employment (column 6) and at the cost of not pursuing host-country education (column 3). We also note that the estimated coefficients on the squared number of victimisation events are close to zero and not statistically significant at any conventional level. Adding polynomials that allow for a more flexible curvilinear relation between victimisation events and integration outcomes are therefore likely to be unnecessary. Similar to our main results, we find no effect of financial victimisation events on economic integration outcomes in Germany.

In table 2.11 we then turn to the results where the different number of victimisation events enter as categorical variables, against the base category of zero victimisation events.

	LFP & Educ.	LFP	Educ.	All employ.	FT employ.	Part-Time or marginal
	(1)	(2)	(3)	(4)	(5)	(6)
1 P.V.E.	0.0443**	0.0623***	-0.0230*	0.0558***	0.0111	0.0446***
	(0.0209)	(0.0212)	(0.0138)	(0.0210)	(0.0163)	(0.0165)
2 or more P.V.E.	0.0335	0.0510*	-0.0360*	0.0082	0.0179	-0.0096
	(0.0283)	(0.0285)	(0.0200)	(0.0308)	(0.0252)	(0.0227)
1 F.V.E.	0.0009	-0.0054	0.0112	0.0062	0.0031	0.0031
	(0.0204)	(0.0210)	(0.0146)	(0.0212)	(0.0166)	(0.0165)
2 or more F.V.E.	-0.0316	-0.0226	-0.0223	0.0066	0.0110	-0.0044
	(0.0259)	(0.0262)	(0.0162)	(0.0270)	(0.0220)	(0.0206)
Observations	2,314	2,314	2,314	2,314	2,314	2,314
R-squared	0.307	0.310	0.210	0.259	0.239	0.155
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals in employment/labour force participation/education, zero otherwise. The acronym P.V.E. denotes physical victimisation events(s) and F.V.E. financial victimisation events(s). LFP means labour force participation. Educ. means education. FT means full-time. Results are only shown for our preferred specification, corresponding to column (2) in the main result tables. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C.origin is the country of origin.

Table 2.11: Number of victimisation experiences (discrete)

The results show that the main results are driven by individuals in both categories, those that experienced one and those that experienced multiple victimisation events, with no clear pattern emerging. The less precisely estimated coefficients on the multiple victimisation event category are likely a result of the smaller number of observations in this group. Overall, we interpret the results of these alternative victimisation specifications as a confirmation of our main results and the modelling choice of victimisation as a binary indicator.

2.6.2 Effects by victimisation event

Our analyses so far did not differentiate between different types of victimisation events individuals endured on their flight to Germany beyond a distinction between physical and

financial harm. We therefore implicitly assumed that all physical and financial victimisation events have a similar effect on economic integration outcomes. In this subsection, we relax this specification and break down the binary physical and financial victimisation indicators into their respective events. Table 2.12 shows the regression results for the same economic integration measures as in all previous analyses, based on our preferred specification.

	LFP & Educ.	LFP	Educ.	All employ.	FT employ.	Part-Time or marginal
	(1)	(2)	(3)	(4)	(5)	(6)
Exp. robbery	-0.0251	-0.0112	-0.0299*	0.0180	-0.0038	0.0219
	(0.0274)	(0.0275)	(0.0177)	(0.0291)	(0.0226)	(0.0237)
Exp. extortion	-0.0231	-0.0130	0.0088	-0.0359	-0.0385*	-0.0137
	(0.0249)	(0.0253)	(0.0175)	(0.0254)	(0.0202)	(0.0192)
Exp. fraud	-0.0017	-0.0104	0.0000	0.0182	0.0263	-0.0081
	(0.0202)	(0.0206)	(0.0143)	(0.0249)	(0.0202)	(0.0192)
Exp. sexual harass.	0.0180	0.0467	0.0061	0.164**	0.0110	0.153**
	(0.0765)	(0.0777)	(0.0507)	(0.0701)	(0.0471)	(0.0633)
Exp. shipwreck	0.0083	0.0231	-0.0123	-0.0131	-0.0028	-0.0103
	(0.0254)	(0.0256)	(0.0157)	(0.0255)	(0.0191)	(0.0196)
Exp. physical attack	0.0611**	0.0670***	-0.0346**	0.0488	0.0399*	0.0089
	(0.0248)	(0.0250)	(0.0177)	(0.0302)	(0.0240)	(0.0237)
Exp. incarceration	0.0205	0.0252	-0.0065	0.0117	0.004	0.0077
	(0.0222)	(0.0226)	(0.0154)	(0.0231)	(0.0187)	(0.0177)
Observations	2,314	2,314	2,314	2,314	2,314	2,314
R-squared	0.307	0.310	0.210	0.261	0.242	0.155
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals in employment/labour force participation/education, zero otherwise. FT means full-time. LFP means labour force participation. Educ. means education. Results are only shown for our preferred specification, corresponding to column (2) in the main result tables. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C.origin is the country of origin.

Table 2.12: Breakdown by type of victimisation event

We start by noting that the lack of an association between financial victimisation and our economic integration outcome measures is generally not driven by opposing signs of financial victimisation events that otherwise affect the outcome. Two individually estimated coefficients among the financial victimisation events show an association with outcome measures that is significant at the 10% statistical level. Having experienced robbery during the journey shows a negative association with pursuing host-country education (column 3), while individuals who experienced extortion are moderately less likely to be full-time employed (column 6). Absent other clear patterns, a suggestive explanation for these findings is that, unlike fraud, robbery and extortion may be more likely to occur concomitant with physical victimisation events. However, we also note that statistical significance at the 10% level of two out of 18 coefficients estimated is well within the expected range of randomly drawn coefficients and should therefore be interpreted with

care. On the other hand, the results obtained on the physical victimisation events show a much clearer pattern. Column (2) shows that events of sexual harassment, shipwreck, physical attacks and incarceration are all positively linked to labour force participation, while column (3), with the exception of sexual harassment, shows a negative association of these events with pursuing host-country education. We further note that while the estimated coefficients on the different physical victimisation events in table 2.12 are generally of the same sign, having experienced physical attacks has the strongest effect on economic integration outcomes, both in magnitude and statistical precision. The finding is well in line with the victimisation literature that finds physical abuse to be the strongest predictor of mental well-being [Dolan et al., 2005, Mahuteau and Zhu, 2016, Johnston et al., 2018]. A further suggestive explanation is that physical abuse directly carried out by agents along the migration route leaves larger mental scars on victims than second-order victimisation events such as shipwreck and short-term incarceration that are only indirectly caused by perpetrators.

2.6.3 Heterogeneous effects by major countries of origin

One of the main concerns using sensitive survey data on victimisation is the reliability of responses as discussed in subsection 2.4.5. In our setting, a further relevant problem is that respondents could be inclined to over-report victimisation if they think vulnerability is expected of them by the host community. This general problem of sensitive survey questions is exacerbated by the uncertain legal status many of the surveyed refugees faced at the time of their first interview in 2016 when about 36 percent of the respondents were still awaiting their asylum decision. Two institutional features alleviate this concern. First, interviewers make it clear to all respondents that the survey is conducted independent of the asylum procedure itself and information provided in the survey cannot be used against surveyees. Second, asylum is granted based on individuals' safety in their home country, rather than during their journey. Nevertheless, some respondents may still give answers they deem favourable with regards to their chances of receiving protection. If such misreporting was systematically correlated with future individual labour market outcomes, this could bias the estimated coefficients. We address the issue by splitting up our analysis by country of origin, exploiting the fact that Syrians who were displaced from Syria between 2014 and 2016 are particularly unlikely to give socially desired survey responses. Due to the war in Syria that spread across the entire country, the rate of Syrians who were granted protection in Germany was extremely high and stood at 97 percent over

our observation period. In fact, the German government acknowledged the general need for protection of displaced Syrians and introduced so-called simplified asylum procedures for Syrians already in November 2014. These allowed Syrian asylum seekers to get their asylum status granted by simply filling in a ten-page questionnaire and by proving that they were actually from Syria [Grote, 2018].[10]

Tables 2.13 show the results for the countries of origin where more than 100 respondents are available, grouping Iraq and Iran as well as Afghanistan and Pakistan.

	LFP			Education		
	Syrian	Iraq & Iran	Afghanistan & Pakistan	Syrian	Iraq & Iran	Afghanistan & Pakistan
	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	0.0551**	0.0578	0.0623	-0.0190	-0.0124	-0.130***
	(0.0228)	(0.0503)	(0.0745)	(0.0156)	(0.0267)	(0.0488)
Financial victim.	-0.0170	0.0708	0.0432	-0.00940	0.0382	-0.00181
	(0.0219)	(0.0526)	(0.0685)	(0.0153)	(0.0236)	(0.0378)
Observations	1,543	375	203	1,543	375	203
R-squared	0.272	0.378	0.588	0.182	0.423	0.500
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
	FT employed			Part-Time or marginal		
	Syrian	Iraq & Iran	Afghanistan & Pakistan	Syrian	Iraq & Iran	Afghanistan & Pakistan
	(7)	(8)	(9)	(10)	(11)	(12)
Physical victim.	0.00303	0.0586	-0.0761	0.0319*	0.0481	-0.00439
	(0.0180)	(0.0386)	(0.0662)	(0.0178)	(0.0344)	(0.0614)
Financial victim.	0.0115	0.00810	0.0388	-0.0127	0.0357	0.0741
	(0.0172)	(0.0374)	(0.0522)	(0.0173)	(0.0327)	(0.0525)
Observations	1,543	375	203	1,543	375	203
R-squared	0.162	0.286	0.504	0.089	0.251	0.509
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals in the labour force/education, zero otherwise. FT means full-time. Results are only shown for our preferred specification, corresponding to column (2) in the main result tables. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C. of origin is the country of origin.

Table 2.13: Outcomes by main country of origin

We conclude that the main results of physical victimisation are not driven by individuals of any country of origin in particular. Columns (1), (2) and (3) show that physically victimised individuals are more likely to join the labour force rather than pursuing host-country specific education (columns 4, 5 and 6), regardless of their origin. We note that the magnitude of effects is largest for migrants originating from Afghanistan and Pakistan (columns 3 and 6). A likely explanation can be found in subsection 2.6.2: In our sample, Afghans and Pakistanis were significantly more likely to experience physical abuse (28%) than Syrians (10%) and individuals originating from Iraq or Iran (14%). The association of physical victimisation with part-time employment (column 11, 12 and 13) on the other

[10]While at the time of the policy introduction there was public fear of abuse of these simplified procedures, a later assessment by the German Federal Office for Migration and Refugees found that 99.6 percent of applicants had filled in the questionnaires truthfully and were indeed Syrian Nationals [German Federal Office for Migration and Refugees, 2020].

hand is driven primarily by Syrians, both in magnitude and statistical precision. We note that Syrians constitute by far the largest group and the more precisely estimated results are therefore not unexpected. Financial victimisation shows no association with economic integration measures across the different estimations, adding further robustness to our main results.

2.6.4 Heterogeneous effects by gender

The asylum seekers entering into Germany between 2013 and 2017 mainly originated from countries where women have culturally different economic roles than men [Fuchs et al., 2020]. If individuals regress to a present-oriented mindset in response to victimisation experiences, negative events occurring during the flight to Germany could have effects on the economic integration of refugees that differ between men and women. For example, if joining the labour force represents a bigger step for women than men, potentially traumatic events and their negative effect on mental well-being may discourage women relatively more from becoming economically active. Table 2.14 therefore shows the main results of our preferred specification by gender.

	LFP		Education		FT employ.		Part-Time or marginal	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Physical victim.	0.0422**	0.0641	-0.0195	-0.0351*	0.0268	-0.00236	0.0225	0.0439**
	(0.0210)	(0.0447)	(0.0171)	(0.0183)	(0.0223)	(0.00941)	(0.0204)	(0.0222)
Financial victim.	-0.00164	0.0184	-0.000105	-0.00479	0.00727	0.0136	0.00103	-0.00224
	(0.0206)	(0.0426)	(0.0168)	(0.0194)	(0.0215)	(0.0110)	(0.0192)	(0.0235)
Observations	1,472	734	1,472	734	1,472	734	1,472	734
R-squared	0.261	0.319	0.227	0.300	0.244	0.249	0.171	0.185
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals in the labour force/education, zero otherwise. FT means full-time. Results are only shown for our preferred specification, corresponding to column (2) in the main result tables. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C. of origin is the country of origin.

Table 2.14: Outcomes by Gender

The results indicate that both men and women are affected very similarly by victimisation events. While noting that the smaller samples lead to a loss in statistical precision, the estimated coefficients on physical victimisation in columns (2) and (4) indicate that physical victimisation events affect women's decision to join the labour force instead of pursuing host-country education at an even larger magnitude than men's (column 1 and 3). For both men and women, the higher probability to join the labour force following physical victimisation only increases the uptake of part-time employment (columns 7

and 8). Financial victimisation shows no association with economic integration outcomes when splitting the sample between men and women. We therefore conclude that our main results are not driven by any gender in particular.

2.6.5 Heterogeneous effects by arrival cohort

One of the main challenges in the setting at hand is to separate selection effects at different expected victimisation (and thus, risk) levels from individual level effects of victimisation. In our empirical strategy laid out in section 2.4, we discuss our approach in detail and argue that narrowly defined fixed effects relating to both the time and geography of migration (and their interaction) are necessary to mitigate the issue. A related issue is that asylum seeker arrival cohorts may differ significantly in their composition over time for entirely exogenous reasons: Displacement happens in different geographical regions at different points in time, leading to heterogeneous arrival cohorts that reach their destination when labour market conditions are potentially more or less favourable to integration. If, for example, arrival cohorts differ along their educational attainment and also face a different victimisation risk, this could provide an explanation for between-cohort differences in labour market outcomes. While our empirical specification discussed in section 2.4.3 covers this issue by including dyadic fixed effects that limit variation to within-cohorts, it is worth exploring if our results are driven by any arrival cohort specifically. To do so, we split the sample by the different arrival cohorts. The regression results for three different cohorts arriving in Germany in 2013-2014, 2015 and 2016-2017 are shown in table 2.15.

	LFP			Education		
	2013-2014	2015	2016-2017	2013-2014	2015	2016-2017
	(1)	(2)	(3)	(4)	(5)	(6)
Physical trauma	0.0338	0.0487**	0.157**	-0.0295	-0.0318**	-0.0196
	(0.0567)	(0.0229)	(0.0751)	(0.0419)	(0.0152)	(0.0458)
Financial trauma	-0.0515	0.00365	0.0375	0.000815	0.0110	-0.0440
	(0.0507)	(0.0226)	(0.0722)	(0.0453)	(0.0145)	(0.0468)
Observations	408	1,527	292	408	1,527	292
R-squared	0.526	0.299	0.463	0.425	0.190	0.321
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
	FT employ.			Part-Time or marginal		
	2013-2014	2015	2016-2017	2013-2014	2015	2016-2017
	(7)	(8)	(9)	(10)	(11)	(12)
Physical victim.	-0.0135	0.0003	0.0202	0.0510	0.0274	0.0339
	(0.0536)	(0.0176)	(0.0378)	(0.0666)	(0.0172)	(0.0547)
Financial victim.	-0.0557	0.0276	0.0345	-0.0871	-0.0118	0.107*
	(0.0497)	(0.0179)	(0.0294)	(0.0576)	(0.0164)	(0.0552)
Observations	408	1,527	292	408	1,527	292
R-squared	0.434	0.237	0.265	0.490	0.136	0.366
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals in the labour force/education, zero otherwise. Results are only shown for our preferred specification, corresponding to column (2) in the main result tables. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C. of origin is the country of origin.

Table 2.15: Outcomes by Cohort

No clear pattern emerges when comparing the effect of physical victimisation events on economic outcomes across arrival cohort. All estimated coefficients are of the expected sign. We note that the 2015 cohort is far larger than the 2013-2014 cohort and the 2016-2017, which accurately reflects the magnitude of asylum seeker inflows into Germany but is also likely to explain the higher statistical precision on estimates related to the cohort (columns 2, 5, 8 and 11). We therefore conclude that differences in arrival cohorts are unlikely to provide an explanation for the effect of physical victimisation on refugees' decision to invest into host-country education or take up employment shortly after arrival.

2.7 Testing alternative mechanisms

One of the limitations of the "loss of future directedness" (or higher time discounting rates) channel we propose as an explanation to our findings is that it is not easily distinguishable from more general mental well-being effects. Since we cannot test the "loss of future directedness" directly, we go through alternative mechanisms that could plausibly explain our findings as laid out in section 2.2.

We first zoom into the asylum procedure more closely in section 2.7.1 to analyse if our results could be mechanically driven by design features of the German asylum system. In 2.7.2, we then test if disproportional financial hardship among the physically victimised could explain their faster uptake of low-income employment. Finally, in section 2.7.3, we test if the negative experience during the journey could have an off-putting effect on victimised individuals' intention to stay in Germany shortly after arrival, which could in turn make the investment into host-country specific human capital less attractive. In this section we show that none of this mechanisms is likely to be driving our findings.

2.7.1 Institutional design: Asylum procedures

If asylum procedures take less (more) time for victimised individuals, they will get earlier (later) access to the labour market upon arrival. This would mechanically link victimisation to a faster (slower) labour market integration and potentially bias the results. In our setting, it is conceivable that victimised individuals have a more legitimate claim for protection and their refugee status could therefore allow them to integrate into the labour market in larger numbers by design.[11] We note here that asylum is granted based on reasons related to human rights violations and persecution individuals face in the country of origin (rather than the journey). Yet, asylum decisions are made based on the judgement of asylum officers and the lower mental health of those who were victimised during the journey could make their asylum claims more convincing, leading to faster procedures.

To test this, we compare the outcome and the length of asylum procedures between victimised and non-victimised individuals explicitly in table 2.16. There is no visible difference between financially victimised individuals and those that were not victimised with regard to the share that ultimately received protection status. Among the physically victimised, the share was even slightly lower (69.4 percent) than among the non-victimised (75.4 percent). We also note that the average unconditional duration of the asylum procedure is slightly longer among both physically and financially victimised individuals compared to those who did not experience victimisation during their journey.

[11] We partly address address this point by including a categorical variables capturing each individual's asylum status in our main specification.

Variable	Mean	Std. Dev.	Min.	Max.	N
Non-victimised					
Asylum granted	0.754	0.431	0	1	1258
Length of asylum procedure in months	7.552	6.359	0	47	841
Physically victimised					
Asylum granted	0.694	0.461	0	1	994
Length of asylum procedure in months	7.782	6.372	0	44	595
Financially victimised					
Asylum granted	0.734	0.442	0	1	1026
Length of asylum procedure in months	7.473	6.151	0	44	673

Note: The variable measuring the length of the asylum procedure is not available for all asylum seekers. Our tests that the variable is not systematically missing are available upon request.

Table 2.16: Summary statistics asylum pocedure

The German asylum system has a second key institutional feature that could encourage fast employment among specific segments of the asylum seeking population. Despite options being very limited in scope, obtaining employment before asylum can improve the chances of receiving a temporary protection status ("Duldung") in Germany [Brücker et al., 2019]. Finding employment upon arrival is therefore particularly incentivised for migrants with a low probability of receiving full protection status since employment disproportionally increases the probability of being allowed to stay in Germany for these individuals. If some individuals' migration decision is motivated by economic reasons in addition to humanitarian reasons and these individuals take higher risks during their journey, these asylum seekers could then also be more motivated to increase their chances of being granted a permission to stay by taking up employment before the end of their asylum procedure.

We test this possibility on the IAB employment biography data by mapping employment rates between victimised and non-victimised refugees for a) the time of arrival and the point in time when asylum was granted and b) after asylum was granted. The exercise of a pre-trend and post-trend comparison allows us to test at what point employment rates start to diverge.

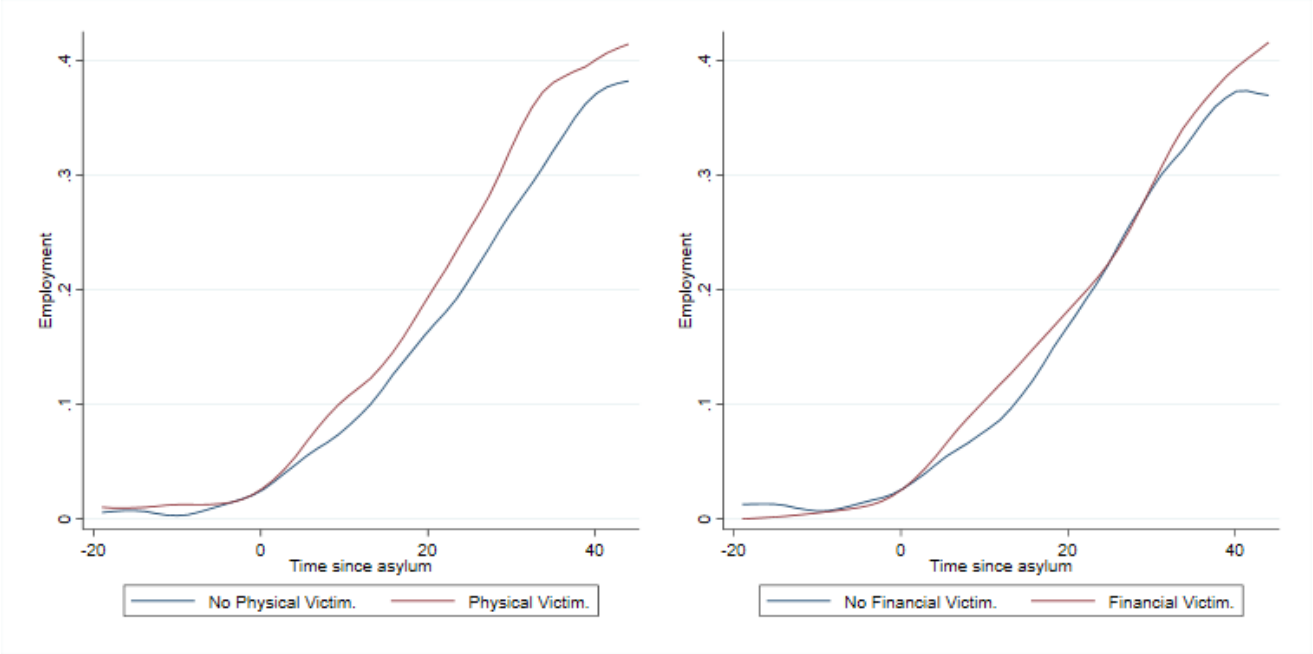


Figure 2.2: Pre- and post-protection trends in employment

Figure 2.2 shows the result of this exercise, with the x-axis starting at the time of arrival and $t = 0$ indicating the month in which asylum was granted. We do not find any evidence that employment rates diverge already prior to the end of the asylum procedure.

2.7.2 Behavioural changes due to financial difficulties

People smugglers have been documented to be responsible for abuse of asylum seekers during their journey and often charge large amounts for their services [Albahari, 2018]. Table 2.22 documents that the average unconditional amount physically and financially victimised asylum seeker paid to escape agents exceeded the amount the non-victimised paid by EUR 1420 and EUR 1802 Euro respectively. An alternative hypothesis to explain our main results is therefore that the faster labour market integration of victimised individuals is caused by the attempt to recover the relatively high cost of the journey quickly upon arrival. We note that, since the reported average group-level differences in the amount paid to people smugglers are unconditional, the opposite effect is also conceivable: Victimisation could reflect a lower ability to pay human smugglers which could have led to retaliatory violent acts by agents who demanded payment.

We partly address this concern in our PDS specification in which we include the incomplete smuggler costs variable as a control. Our data allows us to test the "financial hardship" hypothesis in a second way: We approximate the level of financial precarious-

ness of refugees in Germany by the extent to which survey respondents state to be worried about their personal finances at the time of the first interview. The results of regressing a binary indicator that takes the value 1 for those individuals who state they are "very concerned about their finances" (and 0 otherwise) using our preferred specification is shown in table 2.17.

Very worried about finances	Cross Section		Panel Data	
	Benchmark	Fixed effects	Benchmark	Fixed effects
	(1)	(2)	(3)	(4)
Physical victim.	0.0149	0.00991	0.0295	0.0246
	(0.0217)	(0.0233)	(0.0186)	(0.0210)
Financial victim.	0.0310	0.0413*	0.0332*	0.0450**
	(0.0210)	(0.0221)	(0.0180)	(0.0199)
Observations	2,180	2,180	3,380	3,380
R-squared	0.102	0.194		
C.origin*Departure FE	No	Yes	No	Yes
Baseline Controls	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals who state they are "very concerned about their finances" at the time of the interview. The cross-sectional regressions are run on the first available observation of each individual, 19 months after arrival on average. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C. of origin is the country of origin.

Table 2.17: Very worried about finances

We do not detect an effect of physical victimisation on financial hardship at any conventional statistical level. Unsurprisingly, financially victimised refugees are indeed more likely to voice concern about their financial situation. In our preferred specification of column (2), the magnitude of the effect is an excess 4.1 percentage points over other refugee groups ($p<0.1$).

2.7.3 Intention to remain in Germany

The hypotheses related to the intention to remain in Germany follow directly from the classic human capital investment model for migrants. The model posits that the longer migrants intend to stay, the more they invest into host-country specific education and the less likely they are to take up low-skill employment in the early years after arrival [Cortes, 2004]. There are two ways in which victimisation events individuals experienced during their journey could conceivably be linked to the intended time of stay in Germany. First, the difficult journey could have had a disenchanting effect on the victimised, in particular if violent acts were carried out by agents such as border police that are associated with the host country (or in the case of the EU, the hosting union). This would lead to a

observed negative effect of victimisation on the intention to stay in Germany and could possibly explain why the victimised invest less into education and training in Germany. Second, experiencing victimisation could be understood as a cost by the affected. If the victimised perceive their negative experiences as an investment, they may want to recover these cost by staying in Germany for as long as possible. In this case, the observed effect of victimisation on the intention to stay in Germany would be negative and our main results would underestimate the true effect of the hypothesised "loss of future directedness".

We test these competing hypotheses by analysing differences in refugees' stated intention to stay in Germany permanently upon arrival in table 2.18.

Intention to stay in GER	Cross Section		Panel Data	
	Benchmark	Fixed effects	Benchmark	Fixed effects
	(1)	(2)	(3)	(4)
Physical victim.	-0.0204	-0.0159	-0.0245**	-0.0178
	(0.0125)	(0.0136)	(0.0106)	(0.0126)
Financial victim.	0.00567	0.00424	0.00662	0.00650
	(0.0116)	(0.0128)	(0.00991)	(0.0117)
Observations	2,180	2,180	3,380	3,380
R-squared	0.055	0.126		
C.origin*Departure FE	No	Yes	No	Yes
Baseline Controls	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals who state they are "intend to stay in Germany permanently" at the time of the interview. The cross-sectional regressions are run on the first available observation of each individual, 19 months after arrival on average. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C. of origin is the country of origin.

Table 2.18: Intention to stay in GER

We first note that the cross-sectional regressions using the first available observation after arrival are most informative. The reason is that with time spent in Germany, the intention to stay in Germany permanently may become endogenous to our outcomes of interest. If individuals don't manage to integrate economically, they may have a lower propensity to continue staying in the host country. Our preferred specification of column (2) shows a small, statistically insignificant negative effect of physical victimisation on the likelihood of wanting to stay in Germany. Thus, the results shown in table 2.18 offer only weak support to the hypothesis that our main results can be explained by differences in the intention to stay in Germany between the victimised and the non-victimised.

2.8 Conclusion

One of the key features of humanitarian migration flows from developing into developed regions of the world has been the extreme conditions under which these movements take place. While it is clear that it is exactly the perilous journey itself that limits and deters migration flows, the unintended consequences of such restrictive policies are not yet well understood. In this paper, we analyse how victimisation during asylum seekers' flight to safety, a direct consequence of restrictive migration policies, interacts with labour market outcomes in the destination country. We first show that physical and to a lesser extent financial victimisation events that occurred during the escape journey negatively affect the mental well-being and health among the refugee population in Germany at the time of arrival. In a second step, we then show that physical victimisation in particular distorts trajectories of economic activity in the destination countries: Physically victimised individuals join the labour force faster through the uptake of part-time or marginal employment instead of investing into host-country specific human capital. We do not find a similar effect for financially victimised refugees, suggesting that, in line with the previous victimisation literature, the act of physical victimisation has relatively stronger effects on life trajectories. Financially victimised experience financial hardship in relatively larger numbers when arriving at their destination, but we do not detect a distortion on economic activity in the host-country.

We conceptualise our findings as a "loss of future directedness", a concept closely related to that of "impatience" (e.g. higher time discounting rates) in the economics literature: Events of physical victimisation lead to less forward-looking decision-making. In the framework of the migrant-specific human capital investment model, this can be seen as a distortion to the trade-off refugees face upon arrival to either invest into education to get access to higher quality employment at a later stage or take up lower-quality employment shortly after arrival. These distortions are considerable in magnitude: Three years after arrival, refugees who were physically victimised during their journey are more than five percentage points more likely to have joined the labour force to work in low-income professions and are more than 3 percentage points less likely to pursue host-country specific education or training.

The results strongly suggest that entry restrictions to asylum seekers, while undoubtedly limiting the numbers of new-arrivals, have short and long-run welfare implications for destination countries beyond those brought about by the absolute number of refugees they host. The victimisation events reported by refugees in the surveys match those sys-

tematically measured around the EU's external border, suggesting that at least some of the physical violence inflicted on asylum seekers is directly carried out by border agents [Arsenijević et al., 2017]. Our findings imply that these deterrent measures have consequences for the mental well-being of asylum seekers that extend into their economic integration in the host-country. In this context, we note that one limitation of our study is that the sample only consists of newly arrived refugees. Distortions to the optimal timing of labour market entry are likely to result in larger welfare losses to the host country in the medium to long-term. The magnitude of these losses are in turn linked to other variables, such as return migration patterns among refugees. Revisiting and quantifying the economic cost of victimisation events at a later stage is therefore a promising avenue for future research.

In addition to uncovering potentially costly repercussions of restrictive migration policies for optimal labour market trajectories in the destination, all of our findings also cast doubt on the notion of a swift labour market integration as a success metric for refugees more generally. While clearly useful to judge the efficacy of supportive integration policies, we show that the speed of labour market integration on the aggregate also reflects unintended consequences of policies that serve entirely different purposes. Headline figures on labour force participation and employment among refugees should therefore be cross-referenced with the type of employment individuals engage in, and the potential loss in human capital stemming from the early uptake of low-quality jobs.

Appendix

2.A Summary statistics of main outcomes and conflict intensity measures

Outcome variables related to mental well-being and health indicators are summarised in table 2.19. These refer to the first time we observe the individual in the panel, corresponding to a time spent in Germany of 19 months on average.

Variable	Mean	Std. Dev.	Min.	Max.	N
Life satisfaction before migration (1-10)	7.09	2.839	0	10	2314
Health before migration (1-10)	8.339	2.419	0	10	2314
Current Life Satisfaction (1-10)	7.166	2.125	0	10	2314
Health after migration (1-10)	7.9	2.434	0	10	2314
MCS: Mental Component Scale	48.346	11.298	4.626	73.259	2277
PCS: Physical Component Scale	53.444	9.966	13.487	77.651	2277

Table 2.19: Mental well-being and health indicators

Outcome variables related to the labour market integration of refugees are summarised in table 2.20. These refer to the first time we observe the individual in the panel, corresponding to a time spent in Germany of 31 months on average.

Variable	Mean	Std. Dev.	Min.	Max.	N
Economically active	0.759	0.428	0	1	2314
Labour Force Participation	0.741	0.438	0	1	2314
Education or training	0.078	0.268	0	1	2314
Employed	0.218	0.413	0	1	2314
Full-time employed	0.109	0.312	0	1	2314
Part-time or marginally employed	0.109	0.312	0	1	2314
Net income	892.736	571.574	0	3100	440

Table 2.20: Economic activity indicators

Finally, table 2.21 shows the summary statistics for the regional conflict intensity measures.

Variable	Mean	Std. Dev.	Min.	Max.
No conflict	0.096	0.295	0	1
Low conflict	0.362	0.481	0	1
High conflict	0.542	0.498	0	1
N		2314		

Table 2.21: Conflict intensity

2.B Group-level comparisons of victimised and non-victimised migrants

We exploit the rich set of questions relating to both pre-migration characteristics of refugees and information related to their escape available in the IAB-BAMF-SOEP refugee survey. We utilise this information to explicitly profile refugees who were physically or financially victimised on their journey to Germany vis-à-vis the non-victimised refugee population along seven key dimensions: a) individual-level baseline characteristics, b) countries of origin, c) individual health status, d) psychological characteristics, e) reasons for migrating to Germany, f) financing of the escape, and g) route characteristics. These profiles are presented as descriptive comparisons between victimised and non-victimised groups in table 2.22.

In addition to the interpretation presented in subsection 2.3.4 of the main text, it is worth noting that the 36% of asylum seekers reporting physical victimisation is slightly above the share found by Arsenijević et al. [2017] whose data gathered in Serbia exclude the final stage of the journey to Germany. A further 39% of asylum seekers report financial victimisation on their journey and 46% of asylum seekers did not experience criminal victimisation, implying some overlap between physical and financial victimisation events.

Table 2.22: Summary statistics by victimisation experience - pre-migration and journey

	No victimisation			*Physical victimisation*			*Financial victimisation*			*Delta-1*	*Delta-2*
	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs	Phys. victimised	Financ. victimised
Baseline characteristics											
Age at migration	30.79	10.47	1698	28.39	9.47	1282	29.64	9.99	1300	2.51***	0.56
Female	0.39	0.49	1797	0.29	0.45	1368	0.29	0.46	1381	0.09***	0.08***
Low education	0.56	0.50	1797	0.59	0.49	1368	0.53	0.50	1381	-0.04**	0.05**
Medium education	0.27	0.44	1797	0.26	0.44	1368	0.27	0.44	1381	0.01	0.00
High education	0.17	0.38	1797	0.15	0.36	1368	0.21	0.40	1381	0.04**	-0.05***
Econ situation bm: Below average	0.25	0.43	1720	0.25	0.43	1289	0.21	0.41	1326	-0.02	0.04**
Econ situation bm: Average	0.49	0.50	1720	0.45	0.50	1289	0.44	0.50	1326	0.04*	0.04**
Econ situation bm: Above average	0.26	0.44	1720	0.30	0.46	1289	0.35	0.48	1326	-0.01	-0.08***
Number Of Children	1.97	2.10	1779	1.69	2.12	1355	1.71	1.99	1369	0.26***	0.23**
Married	0.65	0.48	1796	0.57	0.49	1368	0.61	0.49	1381	0.08***	0.02
Good German before migration	0.02	0.15	1797	0.01	0.10	1368	0.01	0.11	1381	0.01*	0.01
Region of origin											
Syria	0.56	0.50	1797	0.51	0.50	1368	0.60	0.49	1381	0.08***	-0.06***
Iraq or Iran	0.17	0.37	1797	0.14	0.35	1368	0.14	0.35	1381	0.02	0.02
Afghanisan or Pakistan	0.08	0.27	1797	0.17	0.38	1368	0.13	0.34	1381	-0.09***	-0.02
Sub-Saharan Africa	0.04	0.19	1797	0.11	0.32	1368	0.08	0.27	1381	-0.08***	-0.02**
Other	0.15	0.36	1797	0.07	0.25	1368	0.05	0.23	1381	0.06***	0.08***
Health status											
Health before migration (1-10)	8.20	2.63	1766	8.03	2.74	1337	8.14	2.59	1354	0.21*	0.04
Psychological measures											
Life satisfaction before migration	6.78	2.99	1769	6.55	3.21	1334	6.85	3.09	1355	0.35***	-0.12
Willingness to take risk	4.43	3.38	1695	4.69	3.43	1307	4.75	3.39	1324	-0.18	-0.29*
Resilience	24.96	3.54	1534	25.02	3.22	1210	25.03	3.30	1247	0.00	-0.02
Reasons for migration											
War or forced recruitment	0.81	0.39	1788	0.86	0.34	1363	0.89	0.31	1375	-0.03**	-0.08***
Persecution and discrimination	0.58	0.49	1788	0.66	0.47	1363	0.71	0.46	1375	-0.05**	-0.12***
Economic reasons	0.44	0.50	1788	0.49	0.50	1363	0.52	0.50	1375	-0.03*	-0.08***
Family and friends	0.22	0.41	1788	0.22	0.42	1363	0.23	0.42	1375	-0.00	-0.01
Other reason	0.12	0.32	1788	0.15	0.36	1363	0.14	0.35	1375	-0.03*	-0.01
No conflict region	0.18	0.38	1514	0.17	0.38	1153	0.14	0.34	1179	-0.02	0.04**
Low conflict region	0.33	0.47	1514	0.34	0.47	1153	0.35	0.48	1179	0.01	-0.02
High conflict region	0.49	0.50	1514	0.49	0.50	1153	0.51	0.50	1179	0.01	-0.02
Financing the escape											
Paid an escape agent	0.67	0.47	1148	0.86	0.35	1020	0.87	0.34	1015	-0.13***	-0.14***
Costs for Escape Agent on Escape in Euro	3174.26	5483.30	1144	4597.98	6180.03	1017	4976.13	6928.30	1014	-776.68**	-1400.38***
Escape funded by credit	0.07	0.25	1797	0.09	0.29	1368	0.10	0.30	1381	-0.02	-0.03**
Escape funded by savings	0.50	0.50	1797	0.49	0.50	1368	0.53	0.50	1381	0.02	-0.04**
Escape funded by selling assets	0.42	0.49	1797	0.50	0.50	1368	0.55	0.50	1381	-0.04*	-0.12***
Escape funded by relatives	0.30	0.46	1797	0.35	0.48	1368	0.34	0.47	1381	-0.05**	-0.03*
Escape funded by friend	0.09	0.29	1797	0.14	0.35	1368	0.12	0.33	1381	-0.04***	-0.02
Journey features											
Duration of journey	34.50	81.28	1755	61.25	102.67	1334	54.88	95.29	1354	-24.60***	-14.70***
Deviation from expected journey duration	-3.66	69.65	1754	9.00	83.19	1334	6.48	83.93	1354	-11.93***	-8.03**
Central Med: Sea route	0.06	0.24	1235	0.17	0.38	965	0.13	0.33	983	-0.11***	-0.03**
Western Med: Sea route	0.01	0.09	1235	0.01	0.07	965	0.01	0.11	983	0.01	-0.01
Eastern Med: Sea route	0.70	0.46	1235	0.71	0.45	965	0.73	0.45	983	0.01	-0.02
Eastern Med: Land route	0.12	0.32	1235	0.13	0.33	965	0.13	0.34	983	-0.01	-0.01
By plane	0.26	0.44	1797	0.19	0.39	1368	0.22	0.41	1381	0.07***	0.02
Arrived alone	0.21	0.41	1795	0.32	0.47	1368	0.28	0.45	1381	-0.11***	-0.04**

The column Delta-1 shows the difference between non-physically victimised and physically victimised individuals. The minuend includes non-victimised and financially victimised migrants. The column Delta-2 shows the difference between non-financially victimised and financially victimised individuals. The minuend includes non-victimised and physically victimised migrants.

2.C Mental and physical health scores

The mental and physical health scores are constructed strictly following Jacobsen et al. [2017], who describe all necessary calculations in detail (p.23-24). A higher score reflects better health. The mental health scale (MCS) is based on the following questions:

- Did you feel in low spirits and melancholy?
- Did you feel calm and balanced?
- Did you feel full of energy?
- Due to psychological or emotional problems, did you achieve less in your work or everyday activities than you actually intended?
- Due to psychological problems or emotional problems, did you perform your work or everyday activities less carefully than usual?
- Due to health or psychological problems, have you been restricted in terms of your social contact to for example friends, acquaintances or relatives?

The response scale for all questions related to the mental scale is 1 (Very often), 2 (Often), 3 (Sometimes), 4 (Almost never), 5 (Never).

The physical health scale (PCS) is based on the following questions:

- How would you describe your current state of health (scale: 1 (Poor) to 5 (Very Well))?
- If you have to climb stairs, i.e. walk up several floors: Does your state of health restrict you (scale: 1 (A lot), 2 (A little), 3 (Not at all))?
- What about other strenuous activities in everyday life, e.g. when you have to lift something heavy or need to be mobile: Does your state of health restrict you a lot, a little or not at all (scale: 1 (A lot), 2 (A little), 3 (Not at all))?
- How often in the last four weeks did you suffer from severe physical pain?
- How often in the last four weeks, due to health problems of a physical nature, did you achieve less in your work or everyday activities than you actually intended?

- How often in the last four weeks, due to health problems of a physical nature, have you been restricted in the type of tasks you can perform in your work or everyday activities?

The response scale for the last three questions is 1 (Very often), 2 (Often), 3 (Sometimes), 4 (Almost never), 5 (Never).

The regression results of the effect of victimisation on the outcomes are shown in tables 2.23 and 2.24.

Mental Scale	Cross Section				Panel Data	
	Benchmark	Fixed effects	Dyadic FE	PDS	Benchmark	Fixed effects
	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	-1.033*	-1.202**	-1.059	-1.641***	-0.975**	-1.082**
	(0.527)	(0.557)	(0.742)	(0.546)	(0.440)	(0.471)
Financial trauma	-1.495***	-1.643***	-2.454***	-0.599	-1.130***	-1.342***
	(0.497)	(0.537)	(0.741)	(0.521)	(0.422)	(0.459)
Observations	2,235	2,225	1,677	2,127	4,814	4,814
R-squared	0.153	0.230	0.307			
C.origin*Departure FE	No	Yes	Yes	Some	No	Yes
R.origin*Departure*Arrival FE	No	No	Yes	No	No	No
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is a mental health score on a scale from 0 to 100 at the time of the interview. The cross-sectional regressions are run on the first available observation of each individual. Results are only shown for our preferred specification, corresponding to column (2) in the main result tables. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. Arrival FE refers to the year-month of arrival in Germany. C. of origin is the country of origin, R. of origin is the wider region of origin as shown in table 2.22.

Table 2.23: Mental Scale

Physical Scale	Cross Section				Panel Data	
	Benchmark	Fixed effects	Dyadic FE	PDS	Benchmark	Fixed effects
	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	-0.850**	-1.194***	-1.056*	-1.203***	-0.909**	-1.264***
	(0.423)	(0.449)	(0.604)	(0.450)	(0.364)	(0.393)
Financial trauma	-0.730*	-0.640	-0.435	-0.531	-0.225	-0.328
	(0.411)	(0.441)	(0.594)	(0.434)	(0.354)	(0.384)
Observations	2,235	2,225	1,677	2,127	4,814	4,814
R-squared	0.258	0.323	0.383			
C.origin*Departure FE	No	Yes	Yes	Some	No	Yes
R.origin*Departure*Arrival FE	No	No	Yes	No	No	No
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes:The dependent variable is a physical health score on a scale from 0 to 100 at the time of the interview. The cross-sectional regressions are run on the first available observation of each individual. Results are only shown for our preferred specification, corresponding to column (2) in the main result tables. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. Arrival FE refers to the v of arrival in Germany. C. of origin is the country of origin, R. of origin is the wider region of origin as shown in table 2.22.

Table 2.24: Physical Scale

2.D Least absolute shrinkage and selection operators

While the main strength of supervised machine learning methods, such as the least absolute shrinkage and selection operators (LASSO) is prediction, they can be used to select control variables to address omitted variable bias when many potential controls are available [Tibshirani, 1996]. These methods also allow us to consider interactions and non-linearities that theory-driven specifications typically omit. Starting with a general model $y_i = x_i'\beta + \epsilon_i$, the LASSO minimization problem can be written as:

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i'\beta)^2 + \lambda\sum_{j=1}^{p} \mid \beta_j \mid, \tag{2.2}$$

with $i = 1, ..., n$ observations and $j = 1, ..., p$ regressors. There are up to $p = dim(\beta)$ potential regressors. Here p can be very large, potentially even $p > n$.

The second term of equation (3) represents the cost of including many regressors. λ is the penalization term.[12] The effect of the penalization is that LASSO sets the $\hat{\beta}_j$s of the variables that contribute little to the model fit to zero.

Belloni et al. [2014] developed a 'post-double selection' (PDS) method, in which the underlying idea is to estimate separate LASSO regressions to find predictors of the selection equation and the outcome equation using 'rigorous' penalization. The final equation then includes the union of the variables chosen as controls from the previous step.

In our setting with two variables of interest, $PhysicalVictim$ (PT) and $FinancialVictim$ (FT), we amend this method to a post-triple selection. The first step in this procedure is to estimate the outcome equation (labour market outcomes) using LASSO, without including PT nor FT:

$$Y_{i,f,c,t} = x_{i,f,c,t}'\beta_j + \epsilon_{i,f,c,t}, \tag{2.3}$$

where we denote the set of LASSO-selected controls by A. The vector $x_{i,f,c,t}$ these controls are selected from include a large set of time constant and time varying individual characteristics, country of origin fixed effects, year-month fixed effects, year-quarter fixed effects, and the interaction between all these variables.

[12] There are three main approaches to choose λ: cross-validation [Chetverikov et al., 2019], 'rigorous' penalization [Belloni et al., 2014] and information criteria (AIC, AICc, BIC or EBIC).

The second step is to estimate the probability of physical victimisation:

$$PT_{i,c,t} = x'_{i,c,t}\delta_j + \epsilon_{i,f,c,t}, \tag{2.4}$$

where we denote the set of LASSO-selected controls by B.

The third step is to estimate the probability of financial victimisation:

$$FT_{i,c,t} = x'_{i,c,t}\eta_j + \epsilon_{i,f,c,t}, \tag{2.5}$$

where we denote the set of LASSO-selected controls by C.

The final step is to use OLS to estimate

$$Y_{i,f,c,t} = \gamma_1 PT_i + \gamma_2 FT_i + w'_{i,f,c,t}\beta_j + \epsilon_{i,f,c,t}, \tag{2.6}$$

where $w_{i,f,c,t}$ is the union of the selected controls from steps 1,2 and 3 (e.g., $w_{i,f,c,t} = A \cup B \cup C$).

Belloni et al. [2014] argue that LASSO can be used to select controls because moderate model selection mistakes of the LASSO do not affect the asymptotic properties of the estimator of the low-dimensional parameters of interest. Hence, modelling the nuisance component of our structural model can be seen as a prediction problem [Andersen et al., 2008].

2.E Testing for misreporting in the survey data

Table 2.25 shows the results of the tests suggested in section 2.4.5. We first regress the willingness to answer the victimisation questions on the employment status reported during the first interview to test for willingness to answer and social desirability bias (lhs). We then regress the employment status on the willingness to respond to journey-related question to test if our subsample of those who were willing to answer victimisation-related questions can be considered representative (rhs).

Willingness to answer/Social desirability	Replied Journey Questions	**Answered journey-related questions**	FT employ	PT or marg. employ
	(1)		(2)	(3)
Employed 1st interview	0.0159	Replied journey questions	0.0141	-0.0145
	(0.0251)		(0.0105)	(0.0105)
Observations	4,007	Observations	4,007	4,007
R-squared	0.259	R-squared	0.216	0.131
C. of origin x Departure FE	Yes	C. of origin x Departure FE	Yes	Yes
Baseline Controls	Yes	Baseline Controls	Yes	Yes

Huber-White Standard Errors; *p<.1; **p<.05; ***p<.01

Notes: All estimated on the whole sample available. In column (1), the dependent variable is a binary indicator taking the value 1 if the individual agreed to give information on negative experiences during the journey in their first interview, conducted 19 months after arrival on average. In columns (2) and (3), the dependent variable is a binary indicator, taking the value 1 if the respondent engaged in full-time employment and part-time or marginal employment respectively in the month prior to the last interview, conducted 31 months after arrival on average. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C. of origin is the country of origin.

Table 2.25: Willing to answer journey questions

None of the estimated coefficients is significantly different from zero at any conventional level, leading us to accept the null hypotheses of no detected bias.

2.F Main results using alternative specifications

In this section, we present the results shown in the main result section 2.5.1 using alternative specifications.

These specifications differ from our baseline specification shown in equation 2.1 in the following way:

- No controls: Only includes the two victimisation variables of interest.
- Basic controls: Additionally to the two victimisation variables of interest, this specification includes country of origin by year and month left origin fixed effects, arrival cohort fixed effects, migration route fixed effects, federal state of residency fixed effects and months since migrations and its squared term (continuous).
- Fewer controls: Preferred specification, excluding willingness to take risk, mental resilience and the satisfaction with living arrangement.
- Extra controls: Most undocumented migrants before the 2015 refugee crisis accumulated debt with smugglers to be able to finance their journey. While our data indicates that this is not the case for the more recent waves of asylum seekers (which constitute the majority of our sample), we explicitly control for debt with smugglers in this specification for the subsample of individuals for whom this information is available. We do so by including a variable that equals 1 if the individual used a smuggler and 2 if he or she chose not to report whether he or she used a smuggler or not. We further include a variable that equals 1 if the individual financed his/her trip through credit or borrowing. These variables are also considered in the post double-selection LASSO (PDS) regressions.

The outcomes are shown in tables 2.26 and 2.27. The results all point into the same direction when compared to tables 2.3 and 2.4 in section 2.5.1.

Panel A: Life Satisfaction	No controls	Basic controls	Fewer controls	Extra controls
	(1)	(2)	(3)	(4)
Physical victim	-0.182*	-0.205*	-0.191*	-0.157
	(0.104)	(0.114)	(0.113)	(0.0982)
Financial victim	-0.255	-0.270**	-0.195*	-0.0961
	(.101)	(0.111)	(0.110)	(0.0991)
Used Smuggler				-0.0599
				(0.141)
No smuggler information				0.147
				(0.161)
Escape funded by credit/borrow				0.0608
				(0.159)
R-squared	0.006	0.131	0.184	0.352
Panel B: Health Satisfaction	(1)	(2)	(3)	(4)
Physical victim	-0.117	-0.250*	-0.296**	-0.266**
	(0.119)	(0.131)	(0.119)	(0.118)
Financial victim	-0.224*	-0.271**	-0.220*	-0.173
	(0.117)	(0.128)	(0.116)	(0.115)
Used Smuggler				-0.00407
				(0.176)
No smuggler information				0.171
				(0.195)
Escape funded by credit/borrow				-0.131
				(0.188)
R-squared	0.003	0.134	0.302	0.332
Observations	2,261	2,261	2,261	2,261
C.origin*Departure FE	No	Yes	Yes	Yes
Baseline Controls	No	Some	Some	Yes

Huber-White Standard Errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable captures self-reported life satisfaction or self-reported health on a scale from 1 to 10, 10 being the highest. The sample consists of the first observation of each individual in the sample, corresponding to the date of the first interview conducted, 19 months after arrival on average. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. Arrival FE refers to the year-month of arrival in Germany. C.origin is the country of origin.

Table 2.26: Health and Life Satisfaction After Migration - alternative specifications

Panel A: LFP or Education	No controls	Basic controls	Fewer controls	Extra controls
	(1)	(2)	(3)	(4)
Physical victim	0.0497***	0.0461**	0.0380**	0.0373*
	(0.0191)	(0.0203)	(0.0190)	(0.0190)
Financial victim	0.0267	0.0171	-0.00792	-0.00799
	(0.0190)	(0.0196)	(0.0184)	(0.0184)
Used Smuggler				-0.00991
				(0.0301)
No smuggler information				-0.0381
				(0.0326)
Escape funded by credit/borrow				-0.0180
				(0.0280)
R-squared	0.005	0.180	0.309	0.312
Panel B: Labour Force Participation	(1)	(2)	(3)	(4)
Physical victim	0.0662***	0.0641***	0.0574***	0.0562***
	(0.0195)	(0.0206)	(0.0193)	(0.0193)
Financial victim	0.0302	0.0165	-0.00934	-0.00944
	(0.0194)	(0.0201)	(0.0188)	(0.0188)
Used Smuggler				-0.0134
				(0.0305)
No smuggler information				-0.0486
				(0.0334)
Escape funded by credit/borrow				-0.0138
				(0.0284)
R-squared	0.008	0.179	0.313	0.316
Panel C: Education or training	(1)	(2)	(3)	(4)
Physical victim.	-0.0254**	-0.0262**	-0.0295**	-0.0286**
	(0.0116)	(0.0130)	(0.0128)	(0.0128)
Financial victim	0.00542	0.00364	-0.000177	-0.000155
	(0.0119)	(0.0128)	(0.0125)	(0.0126)
Used Smuggler				0.00573
				(0.0201)
No smuggler information				0.0172
				(0.0226)
Escape funded by credit/borrow				-0.0143
				(0.0173)
R-squared	0.002	0.136	0.206	0.210
Observations	2,314	2,314	2,314	2,314
C.origin*Departure FE	No	Yes	Yes	Yes
Baseline Controls	No	Some	Some	Yes

Huber-White Standard Errors; *p<.1; **p<.05; ***p<.01

Notes: The dependent variable is binary and takes the value 1 for individuals in the labour force or in education (Panel A), in the labour force (Panel B) or pursuing host-country education or training (Panel C). The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C. of origin is the country of origin.

Table 2.27: Economic Activity - alternative specifications

2.G Employment outcomes - in labour force

Table 2.28 shows the employment results of table 2.5 estimated only on the sample of those in the labour force.

	Any employment		Full-time		Part-time or marginal	
	Cross S.	Panel D.	Cross S.	Panel D.	Cross S.	Panel D.
	(1)	(2)	(3)	(4)	(5)	(6)
Physical victim.	0.0350	0.0321*	0.00634	0.00814	0.0286	0.0240
	(0.0246)	(0.0190)	(0.0194)	(0.0149)	(0.0194)	(0.0149)
Financial victim.	0.0120	0.00372	0.0143	0.00721	-0.00221	-0.00334
	(0.0245)	(0.0180)	(0.0193)	(0.0140)	(0.0192)	(0.0144)
Observations	1,679	3,344	1,679	3,344	1,679	3,344
R-squared	0.253		0.251		0.167	
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

The dependent variable is binary and captures all employed individuals inside the labour force, with regressions showing employment outcomes for any employment ((1) and (2)), full-time employment ((3) and (4)) and part-time or marginal employment ((5) and (6)). All other types of employment are set to zero in regressions (3)-(6). Columns (1), (3) and (5) show the results of a cross-sectional regression on the last interview conducted. Columns (2), (4) and (6) are estimated on the panel using random effects. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country.

Table 2.28: Employment - labour force only

2.H Full Results

Table 2.29 shows the main results as in section 2.5.1 showing the coefficients on all covariates included, estimated using our preferred specification.

Table 2.29: Fixed Effects Results, main outcomes

	Life Sat.	Health	LFP& Educ	LFP	Educ	All emp.	FT emp.	PT,M emp.	Inc
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Physical victim.	-0.170*	-0.281**	0.0384**	0.0579***	-0.0296**	0.0435**	0.0132	0.0303**	-0.213
	(0.0978)	(0.117)	(0.0189)	(0.0192)	(0.0128)	(0.0190)	(0.0148)	(0.0147)	(0.144)
Financial victim.	-0.105	-0.181	-0.0076	-0.0089	-0.0006	0.0027	0.0059	-0.0032	0.0245
	(0.0978)	(0.114)	(0.0183)	(0.0188)	(0.0125)	(0.0185)	(0.0145)	(0.0142)	(0.145)
Central Med. Route	-0.502**	0.0456	-0.0061	-0.0032	0.0120	-0.0266	0.0039	-0.0305	0.0591
	(0.255)	(0.283)	(0.0473)	(0.0473)	(0.0375)	(0.0531)	(0.0445)	(0.0398)	(0.289)
Western Med. Route	0.124	0.217	0.192**	0.197**	-0.0626*	-0.0296	0.0626	-0.0922**	0.119
	(0.516)	(0.692)	(0.0803)	(0.0828)	(0.0349)	(0.117)	(0.118)	(0.0439)	(0.750)
Eastern Med. Route (land)	0.302*	0.458**	0.0361	0.0418	-0.0351*	-0.0183	0.0076	-0.026	-0.109
	(0.176)	(0.184)	(0.0307)	(0.0314)	(0.0207)	(0.0336)	(0.0274)	(0.0260)	(0.223)
Eastern Land Borders	-0.140	0.897	0.164*	0.157	-0.0253	-0.0024	0.0171	-0.0195	-1.099
	(0.437)	(0.768)	(0.0989)	(0.0988)	(0.0528)	(0.104)	(0.0675)	(0.109)	(1.004)
Plane directly to Ger.	-0.108	-0.199	0.0338	0.0363	0.0699*	-0.0356	-0.0368	0.0012	0.411
	(0.213)	(0.290)	(0.0465)	(0.0479)	(0.0374)	(0.0489)	(0.0300)	(0.0418)	(0.380)
No route info.	-0.0130	-0.0669	-0.0064	-0.0170	0.0197	-0.0428	-0.0057	-0.0371*	0.185
	(0.122)	(0.140)	(0.0289)	(0.0292)	(0.0187)	(0.0274)	(0.0221)	(0.0203)	(0.211)
Female	0.0836	-0.697***	-0.219***	-0.215***	-0.0194	-0.139***	-0.0941***	-0.0447**	-0.700**
	(0.111)	(0.128)	(0.0238)	(0.0246)	(0.0158)	(0.0210)	(0.0146)	(0.0175)	(0.339)
Age	-0.0255	-0.0269	0.0182***	0.0254***	-0.0154***	0.0223***	0.0154***	0.0069*	0.0458
	(0.0314)	(0.0361)	(0.0066)	(0.0067)	(0.0039)	(0.0051)	(0.0038)	(0.004)	(0.0615)
Age squared	0.0002	-0.0003	-0.0003***	-0.0003***	0.0002***	-0.0003***	-0.0002***	-0.0001**	-0.0005
	(0.0004)	(0.0005)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0009)
Low conflict intensity	0.477	-0.902***	0.104	0.104	0.0147	0.0701	0.0919**	-0.0218	-0.125
	(0.410)	(0.310)	(0.0765)	(0.0758)	(0.0257)	(0.0611)	(0.0356)	(0.0539)	(0.590)
High conflict intensity	0.495	-0.635**	0.0896	0.0762	0.0259	0.0541	0.0888**	-0.0348	-0.0760
	(0.404)	(0.304)	(0.0760)	(0.0753)	(0.0251)	(0.0603)	(0.0348)	(0.0534)	(0.557)
Willingness take risk	-0.0052	0.0167	0.0003	-0.0012	0.0043**	0.0013	-0.0021	0.0033	-0.0303*
	(0.0148)	(0.0167)	(0.0027)	(0.0027)	(0.0018)	(0.0027)	(0.0021)	(0.0022)	(0.0171)
Resilience	0.0650***	0.0366**	0.0067**	0.0058**	0.0031*	-0.0025	0.0003	-0.0028	0.0113
	(0.0152)	(0.0162)	(0.0028)	(0.0028)	(0.0016)	(0.0027)	(0.0021)	(0.0021)	(0.0199)
Life satisf. BFM	0.0337	-0.0390	-0.0054	-0.0031	0.0003	-0.0021	0.0015	-0.0037	0.0216
	(0.0235)	(0.0252)	(0.0037)	(0.0038)	(0.0024)	(0.0038)	(0.003)	(0.0032)	(0.0253)
Good German BFM	-0.262	-0.327	-0.0500	-0.0184	-0.0424	0.138*	0.0365	0.102	-0.260
	(0.480)	(0.466)	(0.0713)	(0.0692)	(0.0431)	(0.0831)	(0.0521)	(0.0815)	(0.637)
No info. on German BFM	-1.438*	0.392	0.0991	0.143	-0.0807	-0.0645	-0.0248	-0.0397	-0.172
	(0.810)	(0.514)	(0.163)	(0.167)	(0.0561)	(0.0861)	(0.0799)	(0.0424)	(1.689)
Help to move	0.149	0.207*	-0.0670***	-0.0610**	-0.0177	-0.0145	-0.0166	0.0021	-0.153
	(0.115)	(0.123)	(0.0236)	(0.0239)	(0.0140)	(0.0209)	(0.0153)	(0.0168)	(0.183)
Arrived alone	0.208*	-0.254*	-0.0159	-0.0008	0.0012	0.0097	0.0075	0.0023	-0.0489
	(0.124)	(0.137)	(0.0206)	(0.0211)	(0.0177)	(0.0243)	(0.0195)	(0.0189)	(0.158)
Employed before mig.	-0.0732	-0.0765	0.112***	0.137***	-0.0181	0.0469**	0.0329**	0.0141	-0.0853
	(0.121)	(0.128)	(0.0253)	(0.0262)	(0.0176)	(0.0213)	(0.0156)	(0.0172)	(0.228)

... continue

... continue

Upper Second. Educ.	-0.0707	0.177	0.0413*	0.0456*	-0.0075	0.0124	0.0106	0.00182	-0.0940
	(0.118)	(0.133)	(0.0233)	(0.0238)	(0.0154)	(0.0230)	(0.0183)	(0.0177)	(0.167)
Post-Sec. and Uni.	-0.355***	0.0478	0.0239	0.0282	0.0211	0.0405*	0.0156	0.0250	-0.0231
	(0.111)	(0.123)	(0.0213)	(0.0217)	(0.0154)	(0.0218)	(0.0175)	(0.0172)	(0.164)
Time in GER (in months)	-0.0115	-0.0091	0.0184***	0.0171***	0.0080***	0.0028	0.0016	0.0012	0.0235
	(0.0238)	(0.0314)	(0.0036)	(0.0037)	(0.0024)	(0.0035)	(0.0028)	(0.0029)	(0.0338)
Time in GER (in months) sq.	0.0003	-0.0002	-0.0002***	-0.0001**	-0.0001*	0.0001	0.0001	0.0000	-0.00020
	(0.000481)	(0.000691)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0004)
Satisf.with Living Situation	0.332***	0.151***							
	(0.0189)	(0.0206)							
Health BFM. (1-10)	-0.0197	0.318***	0.0050	0.0031	0.0015	0.0108***	0.0031	0.0077***	-0.0321
	(0.0241)	(0.0315)	(0.0043)	(0.0043)	(0.0024)	(0.0037)	(0.0028)	(0.0029)	(0.0473)
Econ. situation BFM Bellow Avg.	-0.118	0.0180	0.0005	-0.0037	-0.0272*	-0.0268	-0.002	-0.0248	0.0771
	(0.128)	(0.150)	(0.0237)	(0.0242)	(0.0150)	(0.0234)	(0.0176)	(0.0186)	(0.207)
Temporary Susp. of Deportation	-0.833***	-0.892***	-0.0473	-0.0400	-0.0290	-0.0306	-0.0024	-0.0282	0.0091
	(0.268)	(0.282)	(0.0417)	(0.0420)	(0.0216)	(0.0373)	(0.0294)	(0.0279)	(0.421)
Request to Leave Germany	-0.255	-0.133	-0.110**	-0.105**	0.0376	-0.0737	-0.0878**	0.0141	0.213
	(0.378)	(0.350)	(0.0532)	(0.0536)	(0.0313)	(0.0490)	(0.0363)	(0.0392)	(0.676)
Decision Still Open	-0.352**	-0.284*	-0.0621	-0.0582	0.0075	-0.0432	-0.0259	-0.0174	-0.396
	(0.138)	(0.155)	(0.0465)	(0.0471)	(0.0229)	(0.0303)	(0.0211)	(0.0233)	(0.747)
Status Unknown	-0.144	-0.500	-0.0345	-0.0504	-0.00925	-0.0342	-0.0039	-0.0304	-0.0475
	(0.260)	(0.325)	(0.0354)	(0.0369)	(0.0241)	(0.0338)	(0.0269)	(0.0253)	(0.269)
Spouse in same HH	0.427***	-0.0227	-0.0319	-0.0116	-0.0392**	0.0358	0.0107	0.0251	
	(0.158)	(0.182)	(0.0273)	(0.0280)	(0.0173)	(0.0270)	(0.0202)	(0.0220)	
Spouse in diff. HH in GER	0.171	0.341	0.0414	0.0303	0.0178	0.0357	0.00451	0.0312	
	(0.219)	(0.254)	(0.0404)	(0.0417)	(0.0380)	(0.0452)	(0.0361)	(0.0354)	
Spouse Abroad	-0.344*	-0.0640	0.0813**	0.0884**	-0.0275	0.0912**	0.0405	0.0507	
	(0.197)	(0.221)	(0.0341)	(0.0352)	(0.0235)	(0.0392)	(0.0290)	(0.0336)	
All children live same HH	0.104	0.0258	-0.0553*	-0.0563*	-0.0418**	-0.105***	-0.0624***	-0.0429*	
	(0.151)	(0.180)	(0.0284)	(0.0288)	(0.0174)	(0.0285)	(0.0223)	(0.0229)	
Some children not in same HH	-0.181	-0.333	-0.0614*	-0.0519	-0.0423**	-0.0486	-0.0313	-0.0173	
	(0.207)	(0.234)	(0.0337)	(0.0342)	(0.0188)	(0.0378)	(0.0288)	(0.0295)	
Observations	2,261	2,261	2,314	2,314	2,314	2,314	2,314	2,314	408
R-squared	0.351	0.331	0.311	0.315	0.209	0.258	0.242	0.154	0.479
C.origin*Departure FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Huber–White standard errors; *p<.1; **p<.05; ***p<.01

Notes: The table replicates the results shown in tables 2.3 (Panel A and B column (2)), 2.4 (Panel A, B and C, column (2)), and 2.5 (column (1), (3) and (5)) of the main text but also shows all coefficients estimated on the control variables included in these regressions. The term FE indicates fixed effects. The term departure refers to the year-month of forceful displacement from the home country. C. of origin is the country of origin.

Chapter 3

The value of host-country education for the labour market position of refugees: Evidence from Austria

3.1 Introduction

One of the key characteristics of refugee employment across the developed world is the low quality of jobs of those who manage to obtain employment. Refugees predominantly find work in low-quality, low-income occupations regardless of education attained before being forced to flee their country of origin [Bloch, 2008, Ruiz and Vargas-Silva, 2018, Zwysen, 2019, Brell et al., 2020]. The magnitude of this phenomenon is sizeable: Dumont et al. [2016] note that, in the European Union (EU), "a full 60 percent of employed tertiary-educated refugees [...] are overqualified for the jobs they occupy" (p.27).

In this study, I focus on the difference in the value of formal education attained in the country of origin and the host country to explain the phenomenon of low-quality employment among refugees in the developed world. I exploit a unique institutional setting in Austria, a country that hosted around 90,000 Bosnian refugees during and after the 1992-1995 Bosnian war, to assess the medium to long-term importance of formal host-country human capital acquisition for the labour market position of young humanitarian migrants. To identify causal estimates of the long-term impact of host-country education on employment and job quality, I exploit the unfortunate reality of humanitarian migrants who were unable to decide at what age they were forced to migrate. Depending on the age at the time of arrival in Austria, young Bosnians entered the Austrian education system on

an ad-hoc basis: Bosnians aged up to 15 at the time of arrival were immediately obliged to attend compulsory schooling in Austria. Bosnians aged between 16 and 19 at the time of migration, who had almost no chance to complete education beyond compulsory schooling in Bosnia, had access to further education such as the degree granting vocational training within Austria's dual education system. Bosnians above the age of 19 who had finished their education in Bosnia received very little support to get their foreign degrees acknowledged and did not receive vocational training to prepare them for the Austrian labour market. The setting thus allows for comparing medium to long-term labour market outcomes of humanitarian migrants who were aged around these education thresholds at the time of displacement.

The results from the corresponding instrumental variable regressions imply that discouraging humanitarian migrants from pursuing host-country specific education in favour of meeting minimum employability requirements may be short-sighted: Over the observation period from five to 27 years after migration, the age-induced attainment of education in Austria led, on average, to a 10 percentage points lower likelihood to work below educational attainment compared to Bosnians who had attained similar education in Bosnia. Similarly, it led to a 9 percentage points lower likelihood to work in a low-skill profession. The gap in employment quality is largest in the earlier years of stay in Austria and then slowly closes over time, but differences are visible even two decades after Bosnians first entered Austria as refugees. Refugees who did not receive any formal education in Austria still earned 16 percent less hourly wage income more than two decades after arrival. Splitting the sample by gender shows that these results are caused by both males and females in the early years after arrival. In the later years, they are largely driven by female Bosnians who, absent Austrian education, are likely to have permanently reverted to professions traditionally open to them in Bosnia prior to forced migration.

3.1.1 Contribution to the literature

The starting point of this paper is the observation that most refugees employed in developed countries work in low-quality jobs even compared to other migrant groups [Bloch, 2008, Connor, 2010, Damas de Matos and Liebig, 2014, Dumont et al., 2016, Ruiz and Vargas-Silva, 2018, Zwysen, 2019, Brell et al., 2020]. For example, Zwysen [2019] shows that in Europe, refugees are significantly more likely to work involuntarily part-time or in jobs of low social status than economic or family migrants even after ten years of stay. Similarly, Dumont et al. [2016] show that 60 percent of employed tertiary educated

refugees in the EU work in professions that do not in fact require such a high level of education. This number compares to an estimated incidence rate of 30 percent among tertiary-educated non-EU born migrants that came for non-humanitarian reasons. At a high level, these observations can be partly explained by limitations in destination choices leading to less selection based on economic characteristics, uncertainty regarding the duration of stay in the host country and potential mental health issues resulting from the experience of war and persecution [Brell et al., 2020]. On a more practical level, refugees likely face large discounts on their education due to the origin of their degrees: Since displacement happened abruptly, refugees almost exclusively attained their education in geographically and culturally distant countries, and bring no host-country specific human capital with them [Dumont et al., 2016].

Academic literature on the cross-border transferability of educational degrees indeed documents that foreign degrees are not easily transferable across borders, in particular from less developed countries where refugees mostly originate from. The returns to education in the hosting country are larger than for similar schooling attained abroad, and the value of foreign education decreases with cultural distance [Aleksynska and Tritah, 2013, Chiswick and Miller, 2009]. Three reasons may explain these observations. First, in the European context, the quality of education vastly differs between origin countries of humanitarian migrants and more economically developed destination countries [Bonfanti and Xenogiani, 2014, Dumont and Monso, 2007, Hanushek and Woessmann, 2011]. A second closely related reason is the weaker signalling effects of foreign diplomas compared to host-country education in the tradition of Spence [1973]. In the migration context, this phenomenon has been coined "screening hypothesis" by Chiswick and Miller [2009]: Risk averse employers may discount schooling abroad more heavily depending on the perceived distance to host-country education. For humanitarian migrants in developed countries this distance is often large. As a consequence, migrants with foreign degrees are also subject to statistical discrimination, making it relatively more difficult to get access to quality employment [Zwysen, 2019]. Third, on the more practical side, humanitarian migrants often struggle to produce evidence for their past qualifications [Commission, 2016b].

Attaining formal degrees has to be distinguished from two other types of host-country human capital in the general context of migration [Zwysen, 2019]: Host-country language acquisition and naturalisation defined as the acquisition of the host-country nationality. Studies on the importance of host-country language skills for income and wages of immigrants in general date back at least to McManus et al. [1983], who showed that language proficiency explains much of the Hispanic wage differences in the US. Since then, a range

of studies have confirmed the causal relationship between language skills and labour force participation, employment and earnings of immigrants in different countries and migration settings [Chiswick, 1991, Chiswick and Miller, 1995, Dustmann and Van Soest, 2002, Berman et al., 2003, Dustmann and Fabbri, 2003, Bleakley and Chin, 2004, Lochmann et al., 2019]. The literature of the effect of naturalisation on the economic integration of migrants typically exploits (changes in) eligibility thresholds to acquire citizenship. Studies find a large positive effects of naturalisation on employment rates and wages among immigrants, with marginalised populations and women benefitting the most from acquiring host-country citizenship [Gathmann and Keller, 2018, Hainmueller et al., 2019, Govind, 2021]. The mechanisms these studies put forward as potential explanations for their findings include access to public sector employment, certainty regarding the stay in the host country and signalling that overcomes discrimination [Govind, 2021]. It is important to note that formal host-country education is likely to include the benefits of language training, and likely captures some of the signalling value provided by naturalisation.

Some descriptive studies specifically analyse the resulting discount refugees face on their formal education in developed countries. Using data from the 2003 wave of the New Immigrant Survey, Connor [2010] first shows that while refugees have similar employment rates compared to non-humanitarian migrants in the US, they earn less and work in occupations of lower status. His results then show that schooling in the US is more highly associated with working in skilled occupations for refugees than economic migrants when total schooling is controlled for. Damas de Matos and Liebig [2014] use 2008 EU-LFS data and show that in Europe, the reason for migration is highly correlated with the origin of the education attained, with the majority of humanitarian migrants holding foreign diplomas. Conditional on the level of education, employment rates of refugees are similar to other migrants but refugees are significantly more likely to be overqualified for the occupation they work in, even once language skills are controlled for. Their findings thus suggest that refugees face a relatively larger discount on their educational attainment than other migrant groups, a finding confirmed by Dumont et al. [2016] using more recent 2014 EU-LFS data. Again using 2008 EU-LFS data, Zwysen [2019] finds similar results and further shows that qualifications obtained in the host-country or having taken steps to get foreign education recognised are associated with a lower likelihood of working part-time involuntarily or in jobs of low social status among migrants who initially came for humanitarian reasons.

While these studies are certainly informative of the overqualification phenomenon among refugees in the developed world, they cannot determine the value of formal host-

country education against any valid counterfactual. Since most of the previous work is based on backward reported data from a single cross-section, estimates over time necessarily compare different refugee cohorts [Connor, 2010, Damas de Matos and Liebig, 2014, Dumont et al., 2016, Zwysen, 2019]. Pursuing a degree in the host-country could further capture unobservable individual characteristics such as motivation and intelligence or household level characteristics such as parental education, which could lead to a bias in the estimated coefficients on where the education was attained [Willis and Rosen, 1979]. Comparing the value of formal host-country education to education attained in the country of origin across the whole age range exacerbates this issue: Among refugees of more advanced age, the self-selection into host-country education is likely to reflect different characteristics than for young refugees who simply continue their interrupted education in a different country.[1]

This paper is the first to provide causal estimates of the long-term value of formal host-country education vis-à-vis similar education in the country of origin. The Austrian setting fulfils the three main criteria that allow to study the effect of formal host-country education on long-term economic outcomes of humanitarian migrants. First, a large number of young refugees entered into the host country at approximately the same time, such that all migrants migrated into similar labour market conditions and there is sufficient variation between young new arrivals who attained education in Austria and those who had attained similar education in Bosnia. Second, the acquisition of formal host-country human capital was to a large extent exogenously assigned to young humanitarian migrants by their age at the time of arrival, a non-manipulatable parameter. Thus, it allows the researcher to identify causal estimates of formal host-country education vis-à-vis origin country education within an age bandwidth of sufficiently similar treated and untreated units. Finally, out and return-migration (repatriation) rates were sufficiently low and, importantly, did not differ significantly between Bosnian and Austrian degree holders of similar educational attainment over time. The 22 years of Austrian microcensus data utilised for this study allow for testing this important requirement.

While the mechanisms that lead to a discount on foreign degrees are well-documented, the research conducted within this study uncovers an additional factor that likely drives the low-quality employment of refugee populations in the developed world: Cultural legacy factors pertaining to female employment. The results show that host-country education

[1] A tempting solution in a regression framework would be to simply control for the age at the time of forced migration; however, due to the very high correlation between the age at the time of forced migration and the attainment of host-country education, this approach would introduce collinearity and it would no longer be possible to interpret estimated coefficients on these variables.

decreased the likelihood of female Bosnian refugees in Austria to work in low-skill professions and below their educational attainment almost four times more than for male Bosnians over the observation period of five to 27 years after arrival. These differences are driven by the later years of stay in Austria when men managed to overcome their discounted Bosnian education, while women did not. In former Yugoslavia, women traditionally held low-level positions on the labour market, regardless of their educational attainment [Pascall and Manning, 2000, Darville and Reeves, 1992]. In Austria, where the institutional setting discouraged host-country education for those who had finished their education in Bosnia, it is likely that women reverted to jobs traditionally open to them in Bosnia. The finding suggests that the education system can play a decisive role in overcoming cultural barriers in the labour market.

The remainder of this paper is structured as follows. Section 3.2 provides background information on the institutional conditions Bosnian refugees faced when arriving to Austria, with a focus on a comparison between the Austrian and Bosnian education systems. Section 3.3 describes the data used for this study, shows basic summary statistics of the working sample and defines the outcome variables. Section 3.4 lays out the empirical strategy. Section 3.5 shows the main results and section 3.6 discusses the role of Bosnian women as a driver of these. Finally, section 3.7 provides a concluding discussion.

3.2 Institutional setting in Austria at the time of the Bosnian war

The Bosnian war lasted from April 1992 to December 1995 and forced around 1.2 million Bosnians to flee their country [Valenta and Strabac, 2013]. Due to the specific ethnic composition of Bosnia-Herzegovina and their geographical proximity, about half of these humanitarian migrants sought refuge in neighbouring Serbia, Montenegro and Croatia. The other half, mostly consisting of Muslim Bosniaks, fled to Western Europe where the main recipient countries were - similar to more current inflows of asylum seekers - Austria, Germany and Sweden [Valenta and Ramet, 2011]. Depending on estimates, between 86,500 [Valenta and Ramet, 2011] and 95,000 [Bendl, 2014] Bosnians arrived in Austria, making it the main recipient country relative to its population size.[2] For political reasons, Bosnians were granted temporary protection without asylum procedures as soon

[2]Bosnians in Austria never received full refugee status and were always considered de-facto refugees [Tretter, 2000]. For the remainder of this paper, I will use the term refugee to refer to Bosnian de-facto refugees.

as they arrived [Tretter, 2000]. This initial temporary protection status imposed large uncertainty on Bosnian refugees, closely resembling the uncertainty during long-winded asylum procedures and time restrictions put on residence permits today. The right of residence was initially granted until June 1994 and was then continuously extended until July 1998, when most Bosnians gained permanent residence status [Hageboutros, 2017].

3.2.1 Bosnian refugees and their access to education in Austria

The temporary protection status gave Bosnian refugees immediate access to the Austrian education system. Since no specific regulation governed refugees' rights to education, Austrian authorities simply applied existing general laws to young Bosnians [Tretter, 2000]. A simplified overview of the Austrian education system around the time of the Bosnian war is shown in table 3.1. The categories based on the ISCED-97 are consistently reported in the Austrian microcensus throughout all observation years from 1998 to 2019.

Expected age at completion	Type of schools and degrees	ISCED-97	Simplified ISCED categories
15 (entry: age 6)	Compulsory schooling or below	0-2	Low
18	Academic Secondary School - Upper Cycle	3A	Medium
16-19	Dual education (vocational school + apprenticeship); Intermediate Vocational Education	3B	Medium
19	Higher Vocational Education	4	Medium
20+	University Education	5A	High
20+	Other tertiary education (e.g. Post-Secondary VET, Industrial Master College)	5B	High

Notes: Own table based on microcensus Austria, which includes Pre-vocational School ("Polytechnische Schule") in ISCED-97 category 2.

Table 3.1: Education categories in Austria

Depending on their age upon arrival in Austria, the probability of attaining Austrian education therefore differed sharply across age-cohorts. Three general groups of young Bosnians can be distinguished depending on the age at displacement from Bosnia. First, Bosnians within the compulsory schooling threshold of age 15 were immediately integrated into Austrian schools. This cut-off age for compulsory schooling was strictly implemented [Tretter, 2000] .[3] This meant both de jure and de facto that all Bosnians aged below 15 were guaranteed to attain an Austrian degree. Second, those above the age of 15 who had completed compulsory schooling in Bosnia had access to upper secondary schooling. For most Bosnians, this meant entering into the Austrian dual education system which offers various options of vocational training. Some restrictions applied in the early years of arrival as the maximum quota of foreign workers of 8 percent (later 9 percent) included

[3] Kauffmann et al. [2002] present anecdotal evidence of Austrian bureaucratic practices at the time, supporting the case that no exceptions were made for Bosnian children who attempted to enter into schooling that was not deemed appropriate for their age.

employment within degree-granting apprenticeships, an institutional feature discussed in more detail in appendix 3.B. I show in appendix 3.C.3 that these initial limitations did not systematically alter the educational trajectories of young Bosnians compared to the slightly older cohort. The upper secondary schooling also included access to the academic secondary schools ("AHS Oberstufe") which, after graduation, grant students access to tertiary university education. However, like older cohorts in Bosnia and likely further aggravated by language barriers [Kauffmann et al., 2002], the share of Bosnians who pursued university education was very low (see also figure 3.2). Finally, the situation for Bosnian refugees who had finished their education before displacement stood in stark contrast to that of younger arrivals. Bosnians of adult age had access to language courses organised locally but did not receive structured additional support to integrate into the Austrian education system [Tretter, 2000].

One of the challenges when placing foreign degrees into the domestic education system is the comparability of degrees. In the Austrian microcensus, respondents and trained interviewers determine the Austrian degree that most closely corresponds to the foreign degree held by the respondent. The information is then used to assign the respondent to a category within the International Standard Classification of Education System (ISCED). For the purposes of this study, ISCED refers to its 1997 system, which came into existence before the start of the observation period in 1998. Greussing [2016] analyses the reporting of foreign degrees in the Austrian microcensus and shows that misreporting is a concern for immigrants from South and East Europe - where Bosnians constitute the major group - that fall into the secondary education category. Her results suggest that misreporting is otherwise low. These findings are corroborated by the comparison of the education systems in Austria and Bosnia at the time of the Bosnian war: While the education systems were very similar along broader education categories, the types of degrees granted within ISCED categories 3 and 4 offer multiple possible translations within and across these medium levels of education. Table 3.2 shows the pre-war education system in Bosnia. While overall, both the broader high and low education categories very closely resemble the Austrian system shown in table 3.1, drawing the line between ISCED categories 3 and 4 based on basic and advanced vocational education is not unambiguous.

Expected age at completion	Type of schools and degrees	Simplified ISCED categories
15 (entry: age 7)	Compulsory schooling or below	Low
16	Technical Qualification Intermediate Vocational Education (basic diploma)	Medium
17	Intermediate Vocational Education (vocational diploma)	Medium
18	Intermediate Vocational Education (full diploma)	Medium
19	Intermediate Vocational Education (vocational baccalaureate diploma)	Medium
19	Academic secondary school	Medium
20+	University Education	High
20+	Other tertiary education (e.g. Professional Studies following voc. baccalaureate, Advanced Vocational Studies)	High

Notes: Own table based on Georgeoff [1982].

Table 3.2: Education in pre-war Bosnia

For these reasons, the analyses in this paper will make use of simplified ISCED categories - low, medium and high - that can be accurately compared between Austria and Bosnia before, during and after the time of the Bosnian war.

3.3 Data

The analyses in this paper draw on Austrian microcensus data from 1995 to 2019, with the main sample consisting of the years 1998 to 2019 when information on the country where the highest educational degree was attained is available. The Austrian microcensus draws a 1 percent representative sample from the Austrian population and a module is conducted by in-person interviews once every quarter. It is designed as a rotating panel with individuals staying in the panel for five quarters.

For the purposes of this study, Bosnian refugees are identified on the basis of their country of origin reported as "Bosnia and Herzegovina" and the year of migration to Austria. Since the war in Bosnia officially lasted from 6 April 1992 until 14 December 1995, all migrants who arrived in Austria between 1992 and 1995 from Bosnia and Herzegovina are classified as refugees. The majority of these humanitarian migrants arrived in Austria in the first year of the war. The distribution across arrival years is displayed in figure 3.1.

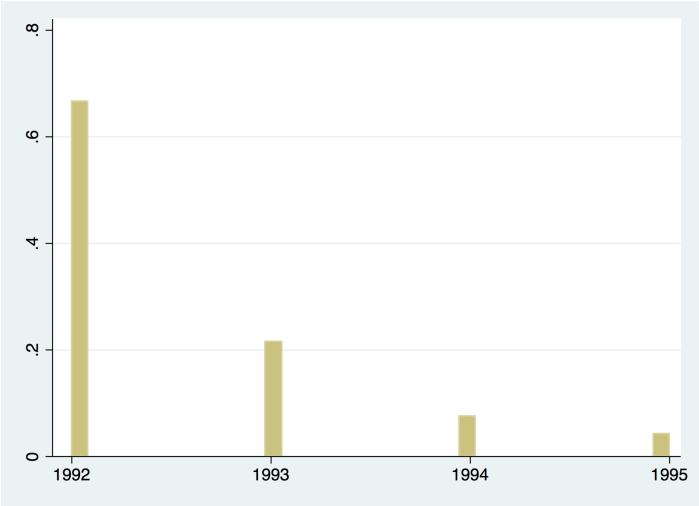


Figure 3.1: Bosnian war refugees by their year of arrival in Austria

Note: N=5283; T=1998-2019; Pooled sample of Austrian microcensus.

Between 1995 and 2003, the questions on the year of arrival and the precise information on the country of origin are only included in the first module of every year (Q1). For consistency, I therefore also only include respondents from the first quarter in the analysis for subsequent years 2004 to 2019. In theory, every surveyed person should be included in Q1 of a year at least once. However, any sample attrition could theoretically limit the sample size this way and quarter 2 (Q2) respondents are therefore included if they do not appear in any Q1. By design, a few individuals appear in two Q1 waves of subsequent years and seven percent of all individuals appear a second time. The Austrian microcensus only includes a question on whether the highest education was attained in Austria or pre-migration from 1998 onwards. All analyses in this study is therefore carried out on the 1998 to 2019 sample. Altogether, the sample consists of 5,559 unique observations over the 22 year observation period.

3.3.1 Working sample

The most striking feature of Bosnian refugees at the time of first arrival in Austria is the young age of migrants, a distribution closely resembling that of humanitarian migrants arriving to the EU in more recent years. More detailed age at the time of forced migration and the education distribution corresponding to the simplified ISCED categories are shown in figure 3.2.

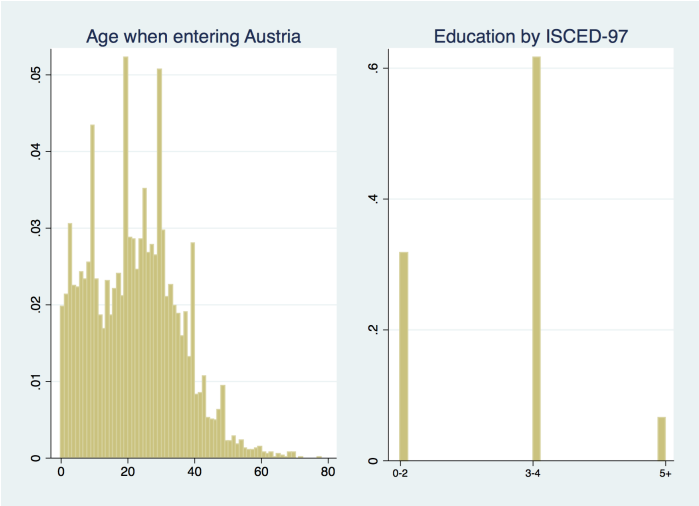


Figure 3.2: Bosnian war refugees by age at the time of migration and education level

Note: N=5283; T=1998-2019; Pooled sample of Austrian microcensus.

Over the pooled sample, the average age of Bosnians migrating to Austria during the Bosnian war stood at 23.1 years. Most Bosnian refugees were educated at a medium level at the time when their respective interview was conducted, corresponding to ISCED categories 3 and 4.

The left skew of the age distribution allows to define a sufficiently large working sample of Bosnians aged 13 to 22 at the time of forced migration to Austria that fulfils two main criteria. First, the bandwidth of the sample does not move too far away from the main education thresholds such that attaining education in Austria is highly correlated with the age at forced migration within the group. Second, the age group falls within the boundaries of adolescence, such that individuals were at a similar phase of psycho-social development when forcibly displaced [Sawyer et al., 2018]. The adolescent age group has been shown to be relatively more resilient to displacement (and potential war) experiences than other age groups [Green et al., 1991]. Subsection 3.4.2 discusses the choice of the working sample in more detail in the context of the instrumental variable strategy. Table 3.3 summarises the main characteristics of the working sample.

Variable	Mean	Std. Dev.	Min.	Max.	N
Low education	0.306	0.461	0	1	1354
Medium education	0.663	0.473	0	1	1354
High education	0.03	0.171	0	1	1354
Age	33.927	6.674	18	48	1354
Age-squared	1195.544	455.627	324	2304	1354
Age at forced migration	17.99	2.923	13	22	1354
Female	0.572	0.495	0	1	1354
Household size	3.701	1.231	1	8	1354
Married (head)	0.747	0.435	0	1	1354
Single (head)	0.208	0.406	0	1	1354
Widowed (head)	0.046	0.209	0	1	1354
Number of children in household	1.792	0.921	0	6	1211
Number of children < 6	0.693	0.825	0	4	1211

Table 3.3: Working sample: Aged 13 to 22 at displacement

One of the caveats of choosing a relatively narrow age at forced migration bandwidth is that the 3.0% share of highly-educated Bosnians among those displaced between 13 and 22 is slightly lower than the sample average of 5.5%. The main reason is that few Bosnians had completed their tertiary education by age 22 before displacement. We turn to the issue of employment quality among the tertiary educated in subsection 3.5.5.

3.3.2 Measuring employment quality

Three different measures of employment quality are calculated to assess the labour market position of Bosnian refugees holding a degree from Austria to those holding a comparable degree from Bosnia. The quality of employment measures that can be consistently produced from the Austrian microcensus between 1998 and 2019 are measures of working below educational attainment and a parsimonious measure of working in a low-skill profession. Information on net hourly wages available in the microcensus between 2011 and 2018 is then used to complement these measures.

The overeducation measure follows the economics of education literature and is calculated using a simplified realized matches procedure in the following way [Chiswick and Miller, 2009, Aleksynska and Tritah, 2013]:

$$Overeducated_{i,n,t} = \begin{cases} 1, if & Education_{i,n,t} > Mode(Education_{n,t}) \\ 0, if & Education_{i,n,t} \leq Mode(Education_{n,t}) \end{cases} \tag{3.1}$$

That is, an individual is overeducated if individual i's education who works in occupation n in year t is above the modal education of the respective occupation, measured based on the simplified ISCED-97 categories, low, medium and high. The Austrian microcensus reports professions based on Austria's mapping of national occupations statistics on the International Standard Classification of Occupation 88 (ISCO-88) at the three digit level. The total number of different occupations reported in the Austrian ISCO-88 is 116. From 2011, the Austrian microcensus switches to ISCO-08 where, at the three digit level, 130 occupations are differentiated. Thus, in a first step, the modal educational level is calculated for each ISCO occupation in every year using the entire sample of employed individuals in the Austrian microcensus. The sample size in the Austrian microcensus is sufficiently large to produce reliable modes. In 1998, the first observation year, the total number of employed individuals stood at 29,833. In 2019, the last observation year, the total number of employed stood at 88,166. In a second step, the education level of each employed Bosnian refugee is then compared to the calculated mode of their respective profession to determine overeducation.

Since the Austrian microcensus only reports education categories rather than years of schooling - and thus the realized matching procedure is based on relatively little variation - a more direct measure of working in a low-skill occupation is also calculated:

$$LowSkill_{n,t} = \begin{cases} 1, if & Mode(Education_{n,t}) \leq ISCED-97, \text{ level } 2 \\ 0, if & Mode(Education_{n,t}) > ISCED-97, \text{ level } 2 \end{cases} \tag{3.2}$$

That is, an occupation is classified as low-skill if the modal education level of workers within the occupation is equal to or below 2 in the ISCED-97 classification or low on the simplified ISCED measure, calculated separately for every survey year t. Thus, all occupations that are primarily carried out by workers with either no education or only compulsory education are classified as low-skill. This measure differs from the overeducation measure in two ways: First, it can be calculated for all individuals, not just those educated above level 2 of the ISCED-97 scale. Second, it relies less heavily on the accuracy of reported education by Bosnian refugees and its translation into the Austrian system. The most frequent occupations Bosnian refugees worked in that the above procedure classifies as low-skill are (i) domestic aid workers, (ii) cleaning staff, (iii) salesperson/supermarket staff, (iv) unskilled labourer in manufacturing and other industries and (iv) construction workers.

From 2011 to 2018, the microcensus data includes information on monthly earnings and monthly hours worked. These information are used to calculate an hourly net wage for each individual to further quantify the long-term value of an Austrian degree vis-à-vis a comparable Bosnian degree.

The main outcome variables are summarised in table 3.4, pertaining to the period average of Bosnians who had spent five to 27 years in Austria.

Variable	Mean	Std. Dev.	Min.	Max.	N	Sample restrictions
Employed	0.815	0.389	0	1	1354	None
Work in low-skill job	0.143	0.35	0	1	1103	Employed only
Work below education	0.112	0.316	0	1	783	Employed only, no low-educated
Hourly net wage	11.89	5.28	2.40	52.26	262	Employed only, years 2011-2018

Table 3.4: Main outcome variables

3.4 Empirical strategy

This section describes the empirical strategy to estimate the value of education attained in Austria compared to similar levels of education attained in Bosnia for labour market outcomes. Subsection 3.4.1 starts with the baseline specification and 3.4.2 then discusses the instrumental variable approach in detail.

3.4.1 Baseline specification

The baseline specification estimated by ordinary least squares (OLS) takes the form

$$Y_{i,t,e} = \beta_t + \kappa_e + \eta EducationInAustria_{i,t} + \zeta IndividualCharacteristics_{i,t} + \epsilon_{i,t}, \quad (3.3)$$

where $Y_{i,t,e}$ is the integration outcome of interest for individual i in education category e at time t. $Y_{i,t,e}$ is a binary variable equal to 1 if individual i is employed, works in a low-skill occupation or works in an occupation he/she is overeducated for respectively. For outcomes related to net hourly wages, $Y_{i,t,e}$ is continuous. $EducationInAustria_{i,t}$ is a binary variable equal to 1 if individual i attained his/her highest education in Austria and 0 if education was attained in Bosnia. Its estimated coefficient is the main coefficient of

interest, η. β_t and κ_e are survey year and (simplified) ISCED-97 education category fixed effects respectively. The inclusion of κ_e is a key feature of 3.3 since it restricts variation to within education categories such that the only difference between individuals is whether that education was attained in Austria or Bosnia. $IndividualCharacteristics_{i,t}$ is a vector of control variables, including a dummy for female Bosnians, the time spent in Austria in years and federal state fixed effects. The latter are included to account for the fact that Bosnian refugees were initially only allowed to take up employment in the federal state they resided in [Tretter, 2000]. The federal state fixed effects also help to control for (the small) differences in the presence of co-ethnic networks and variation in local labour market conditions discussed in appendix 3.A. $\epsilon_{i,t,e}$ is the error term that is clustered on the household level to account for unobserved correlations between household members if members from the same household were interviewed.

Finally, since the quality of employment is only observed for employed individuals, this initial selection could be different for humanitarian migrants who attained education from Bosnia and those who attained education in Austria. For example, in an extreme case where Bosnian degrees are heavily discounted, very few Bosnians without host-country education may find employment. The employed sample of these migrants without host-country education would then differ significantly from the sample of Bosnians who attained host-country education. Appendix 3.C.1 explains the corresponding Heckman selection procedure following Heckman [1979] to adjust the estimated coefficients for the probability of each observation to be included in the sample of the employed and shows that this correction does not alter the obtained results.

3.4.2 Instrumental variable approach

Estimates obtained from equation 3.3 provide an association between holding a diploma from Austria compared to a similar degree from Bosnia. An issue that needs to be addressed to more accurately capture the causal relation between host-country education and the various employment outcomes is the endogeneity of attaining host-country education resulting from omitted variable bias and reverse causality. Pursuing a degree in Austria could capture unobservable individual characteristics such as motivation and intelligence or household level characteristics such as parental education, which would potentially bias the estimated coefficients. The direction of such bias is a priori ambiguous and has occupied scholars for decades [Willis and Rosen, 1979]. In the context of humanitarian migration, it is possible that more capable individuals are able to assess the value

of pursuing host-country education more accurately, or that less capable individuals are discouraged from attaining education due to the higher barriers to entry caused by, for example, the different language. This kind of self-selection could potentially lead to an upward bias in the estimated coefficients. On the other hand, if host-country education is seen as a last resort following an unsuccessful entry into the labour market, we may observe a downward bias in coefficients when estimating the value of host-country education for employment and employment quality. This endogeneity problem of attaining education in Austria is tackled by an instrumental variable approach.

The instrumental variable approach exploits the fact that many Bosnian refugees were of young age when they entered Austria. Since Bosnians were displaced forcibly and unexpectedly, they had no control over their age when migrating. The immediate access to education that was granted to Bosnians and the strictly implemented education laws explained in detail in section 3.2.1 thus led to a strong correlation between the age at the time of arrival and the probability of attaining host-country education (figure 3.3).

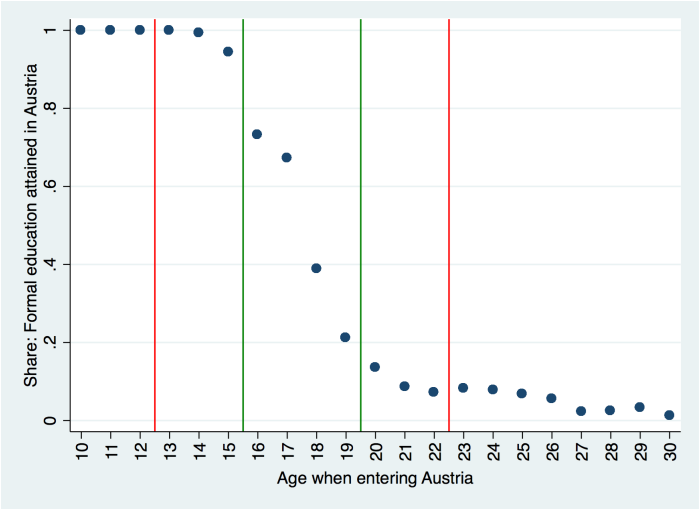


Figure 3.3: Education in Austria and the age at the time of forced displacement

Note: N=2865; T=1998-2019; pooled sample. The green lines indicate education cutoffs. The red lines indicate the bandwidth chosen for the main regression analyses.

In the corresponding instrumental variable regressions, the first stage of the two-stage least squares (2SLS) approach then takes the form:

$$EducationInAustria_{i,t,e} = \beta_t + \kappa_e + \eta AgeAtTimeOfArrival_{i,t} + \zeta IndividualCharacteristics_{i,t} + \nu_{i,t}, \tag{3.4}$$

where $AgeAtTimeOfArrival_{i,t}$ is used to predict the probability of having received education in Austria instead of Bosnia, $EducationInAustria_{i,t,e}$.

The suggested instrumental variable approach laid out above requires a careful reflection. The identifying assumption that age at the time of forced displacement only affects labour market outcome through its effect on education, and thus serves as a valid instrument, likely only holds for an age bandwidth that is sufficiently narrow. Two main rationals guide the choice of the bandwidth. The first rational follows from the literature on displacement studies. Since age is a direct determinant of psycho-social development, the forced migration experience could have had heterogenous effects on migrants across different age brackets, with potential differing repercussions on mental well-being [Porter and Haslam, 2005]. Mental well-being itself could in turn affect the performance on the labour market [Freitas-Monteiro and Ludolph, 2021]. The age group relevant to the research question at hand is that of adolescents, an age group defined by the medical literature to range from 10 to 24 years [Sawyer et al., 2018]. Medical studies on adolescent refugees displaced and hosted in high-income countries are indeed typically conducted on a population aged between 11 and 24 (see Fazel et al. [2012] for a meta-analysis). One of the key reasons for separating adolescents and children from older populations is their relatively higher resilience to stresses of displacement [Green et al., 1991]. Thus, under the assumption of homogeneity in adolescents' psychological response to forced displacement, the bandwidth can then be chosen based on the institutional setting refugees faced in Austria at the time of arrival. The important education cutoffs for Bosnian refugees as shown in table 3.1 were 15, the age when compulsory schooling is completed, and 19, when upper secondary schooling is completed. Thus, Bosnians would typically start upper secondary education at age 16. In our preferred specification, we therefore move, symmetrically, three years to the left and three years to the right of these cutoffs, such that the working sample consists of those aged 13 to 22 at the time of displacement, leaving a total pooled sample of 1354 individuals.

The second related rational for the choice of the bandwidth is technical. For any given bandwidth, estimates are subject to the bias versus precision trade-off faced in regression discontinuity designs [Hahn et al., 2001]. The chosen bandwidth of Bosnians aged 13 to 22 reflects this tradeoff. On the one hand, it leaves a working sample sufficiently large to precisely estimate the effects of interest and study heterogenous treatment effects across time and groups. On the other hand, it does not move too far away from the education cutoffs, avoiding the inclusion of relatively younger and older individuals outside the group of adolescents, where the exclusion restriction is less likely to hold.

Despite these careful considerations, the definition of the bandwidth remains arbitrary. It is therefore important to show that the main results are not dependent on this choice. To do so, I show in appendix 3.C.2 that the main results hold when the bandwidth is widened to any range between 10 to 25 (and thus, to cover the whole spectrum of adolescent age) or narrowed down to 16 to 19, in line with some of the more conservative age band choices of studies on adolescent refugees (see for example Sujoldžić et al. [2006]).

A third consideration related to the validity of the instrument pertains to the external versus the internal margin of schooling attained in Austria. The chosen bandwidth makes the implicit assumption that it is the Austrian degree per se, rather than the internal margin of Austrian schooling measured by the years of schooling in Austria, that predicts employment quality. The institutional setting in Austria makes this assumption likely to hold for all individuals aged above compulsory schooling age: Bosnians who did not finish their upper secondary education in Bosnia were unlikely to get their incomplete upper secondary training acknowledged and had to start over on this part of their education [Tretter, 2000]. For those aged within compulsory schooling age, the preferred working sample of 13 to 22 year olds only includes individuals who received a maximum of three years of compulsory education. Nevertheless, in appendix 3.C.4, I show that it is indeed likely to be the Austrian degree, rather than years of school attendance, that drive the main results in both the working sample and among those who migrated when still being subject to compulsory schooling.

A fourth important requirement of the IV approach is to account for the possibility that forced migration altered education choices in the host country. For example, having attained a level of education above basic compulsory education for those aged slightly above compulsory schooling age in Bosnia when forcibly displaced is unlikely. Those who were displaced at that age could have chosen to pursue higher education in Bosnia but, due to the altered institutional structure, did not do so in Austria. Similarly, incentives to attain education could have been altered for younger cohorts if, for example, young Bosnians or their parents were aware of the value of Austrian education for the job market prospects of their children. Accounting for these potential alterations by conditioning estimates on education category fixed effects is therefore necessary and to a large extend remedies the concern. However, even when variation is restricted to within education categories, the quality of graduates within each education category could be altered compared to a counterfactual of no forced displacement when different individuals select into the different education categories. In appendix 3.C.3, I analyse this potential issue in more detail and show that those aged between 16 and 18 at the time of forced migration

were indeed less likely to be educated beyond compulsory schooling compared to Bosnians slightly older and slightly younger at the time of displacement. The section further shows the results of robustness tests that excludes this age group from the analysis.[4]

The fifth concern pertains to the accurate reporting of age. In theory, if humanitarian migrants knew that a lower age could increase their chances of receiving a residence permit or access to education, manipulation could become desirable. Two observations mitigate this risk. First, the majority of Bosnians entered Austria within the 1965 Austrian-Yugoslav agreement on Visa Policies, which allowed visa-free entry into Austria for three months but required a legal crossing of the Austrian border, suggesting that Austrian authorities had evidence of their age [Franz, 2003]. Second, a conventional manipulation at the threshold test following McCrary [2008] is presented in figure 3.4 for the various relevant age thresholds where these manipulations could have occurred. These tests show no evidence for systematic misreporting of age at the time of arrival.

[4] A very subtle further assumption the suggested approach makes is that there are no differences in how those with a higher innate ability behave at any given age relative to others. For example, if an 18-year old highly capable Bosnian who completed education in Bosnia decided that re-doing their education is worthwhile, while a similarly highly capable 19-year old would not, and all less capable Bosnians aged 18 and 19 would never re-do their education, the age at the time of arrival would partly predict innate ability.

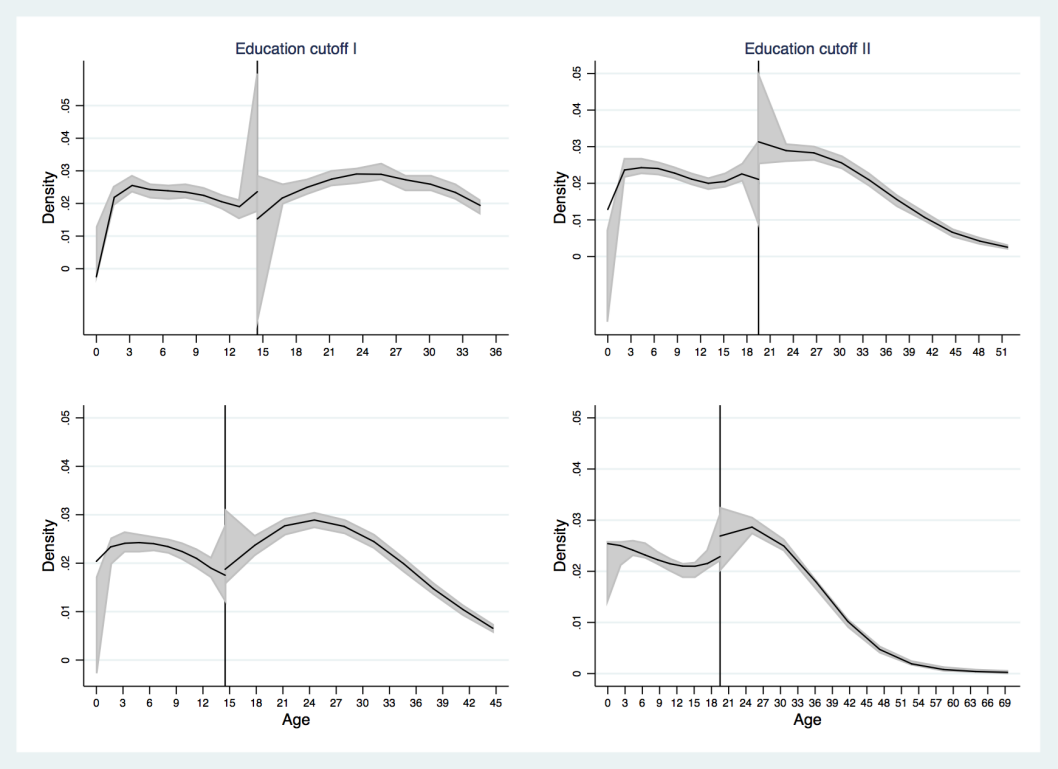


Figure 3.4: McCrary test for manipulation at age thresholds

Note: Education cutoff I refers to the age of arrival between 14 and 15, just before the end of compulsory schooling in Austria. Education cutoff II refers to the age above 19. 19 is the age when all upper secondary education is typically completed and beyond which entering into vocational education may be considered inappropriate by employers. The top graphs show 4th order local polynomials. The bottom graphs show 3rd order local polynomials. The shaded area shows 95% confidence intervals. Age range differs due to data-driven bandwidth selectors.

A sixth requirement for the validity of the instrument is that differences in out and return-migration (repatriation) rates between treated and untreated units must be sufficiently low so the sample composition does not change over time. If sample attrition varied between Bosnians holding an Austrian degree and those who do not due to differing success rates on the labour market, the estimated coefficients on the value of Austrian vis-à-vis Bosnian degrees could be downward biased. Two observations mitigate the concern. First, a key feature of the humanitarian migration flows from Bosnia to Austria during the Bosnian war is that return migration to Bosnia was generally low in its aftermath. A displaced population study carried out by the Bosnian government in 2005 estimated the official number of refugees still present in Austria to be around 82% of those originally displaced [Valenta and Ramet, 2011]. Second, a test comparing the share of Austrian degree holders among Bosnian refugees in the Austrian microcensus over time shows that this share is constant.[5] This is shown in figure 3.5, where the margins are derived from a probit regression of a binary indicator that takes the value one if education was attained in Austria (and zero otherwise) on a categorical survey year fixed effects term.

[5]If, for example, holders of Austrian degrees were less likely to return to Bosnia, this share would increase systematically over time.

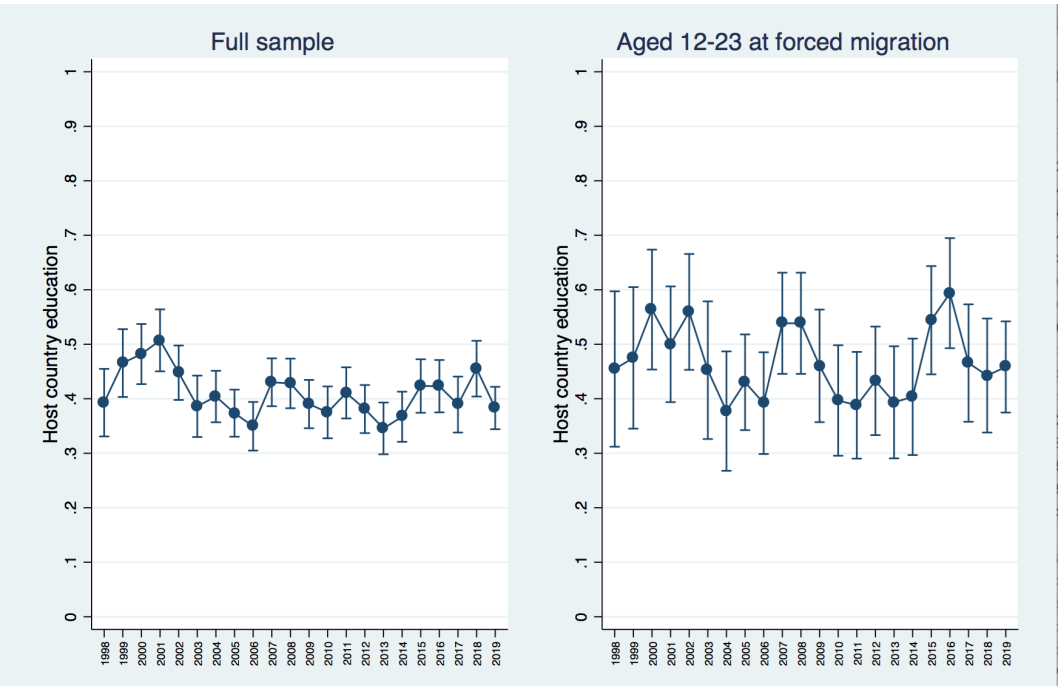


Figure 3.5: Relative attrition rates in the Austrian microcensus by country of educational attainment

Note: Full sample (left hand side) and working sample of those aged 13 to 22 at the time of forced migration (right hand side). The x-axis shows the microcensus survey year. Margins are derived from a probit regression of a binary indicator that takes the value one if education was attained in Austria (and zero otherwise) on a categorical survey year fixed effects term. The vertical bars show 95% confidence intervals.

Finally, we note that in the 2SLS procedure, conditioning estimates on covariates such as age and potential work experience is neither possible nor desirable since these variables would be highly correlated with the instrument by construction.

3.5 Results

This section presents the results. It is structured as follows. Subsection 3.5.1 shows the link between age at the time of arrival in Bosnia and labour market outcomes descriptively. Subsection 3.5.2 shows the main results of a pooled sample regression of the employment and employment quality indicators on the variable indicating whether or not the individual attained formal education in Austria. Subsection 3.5.3 shows how these estimates change over the duration of stay in Austria. Subsection 3.5.4 then shows the effect host-country education vis-à-vis origin country education had on wages after more than two decades of stay in Austria. Finally, subsection 3.5.5 provides a descriptive discussion on the labour market position of the small number of tertiary educated Bosnians that may not be entirely captured by the suggested IV approach.

3.5.1 Descriptive results

Figure 3.6 plots the different outcomes of interest over three brackets of "time spent in Austria" (in years) separately for three groups: Bosnians aged 13 to 15 when being forced to migrate to Austria, Bosnians aged 16 to 19 when being forced to migrate to Austria and Bosnians aged 20 to 22 at the time of migration.

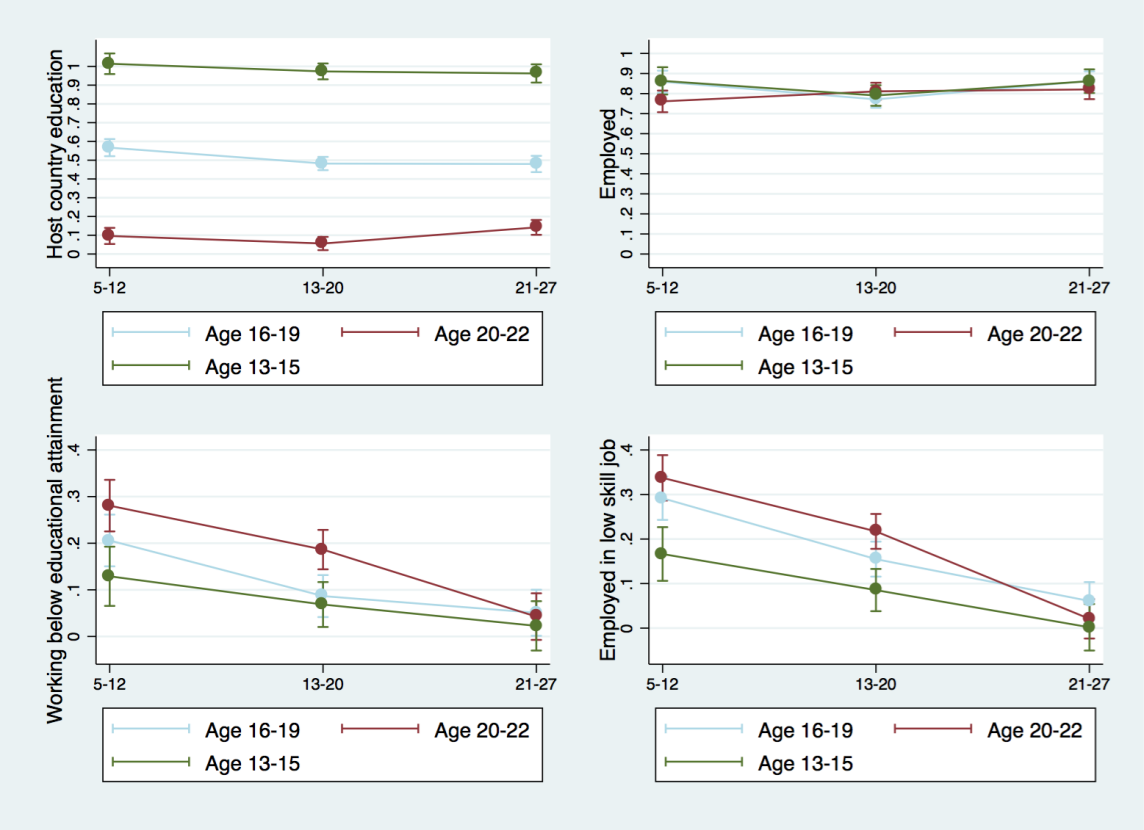


Figure 3.6: Host-country education, employment and employment quality over years of stay in Austria, comparison of age groups

Note: The x-axis displays years since immigration to Austria in three brackets. Age measured when entering Austria as a humanitarian migrant. Margins are derived from a linear regression of a binary outcome indicator on a categorical fixed effects term indicating the "years since migration" brackets, interacted with the "age at the time of arrival" bracket. "Host-country education" is a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. The regression with employment as an outcome is estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. Estimates conditional on gender. Only within-ISCED category variation is considered. The bars show 90% confidence intervals.

The upper left panel shows that over the entire observation period, these age brackets are highly associated with the likelihood of having attained a formal degree in Austria instead of Bosnia. Two observations are striking. First, throughout the observation period from having spent five to 27 years in the country, employment rates are close to indistinguishable between the three groups (upper right panel). Second, the group aged 13 to 15 at arrival performs significantly better on both measures of employment quality than the 20 to 22 year-olds up until the second "time spent in Austria" bracket, when Bosnians had spent between 13 and 20 years in Austria (lower left and lower right panel). After five to twelve years in the country, the conditional share of Bosnians aged 20 to 22 years at arrival who worked below their educational attainment or in a low-skill job stood at 28% and 34% respectively, compared to 13% and 17% among Bosnians aged 13 to 15 at displacement. The 16 to 19 year-olds fall in between. The share of Bosnians in working below educational attainment and working in low-skill jobs then declines gradually over time in all "age at the time of arrival" groups, with shares in age groups that initially experienced a higher incidence of low-quality employment dropping slightly faster, such that the gap becomes small 21 to 27 years after migration.

3.5.2 Main results

We next turn to the regression results when pooling the working sample over the five to 27 year observation period in table 3.5, which displays both the OLS and the 2SLS results of the regression models described in equation 3.3 and 3.4.

	Employed		*Work in low-skill job*		*Work below education*	
	(1)	(2)	(3)	(4)	(5)	(6)
Host-country education	0.045*	0.016	-0.077***	-0.085***	-0.111***	-0.103***
	(0.025)	(0.033)	(0.022)	(0.029)	(0.023)	(0.028)
Time spent in Austria	-0.005	-0.005	-0.002	-0.003	-0.008	-0.008
	(0.015)	(0.015)	(0.011)	(0.010)	(0.012)	(0.011)
Female	-0.041*	-0.045*	0.169***	0.167***	0.106***	0.107***
	(0.024)	(0.024)	(0.020)	(0.020)	(0.022)	(0.022)
N	1354	1354	1103	1103	783	783
R^2	0.048	0.005	0.238	0.083	0.242	0.075
Estimation method	OLS	2SLS	OLS	2SLS	OLS	2SLS
Time FE	Year	Year	Year	Year	Year	Year
Federal state FE	Yes	Yes	Yes	Yes	Yes	Yes
Education category FE	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-test		1743.729		1250.340		1158.215

Standard errors clustered at the household level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: The sample consists of individuals aged between 13 and 22 when arriving in Austria during the war in Bosnia that lasted from 1992 to 1995. The observation period is 1998 to 2019. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. Regressions with employment as an outcome are estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. "Host-country education" is a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia. "Time spent in Austria" is a continuous variable measured in years. The first-stage F-test refers to the Kleibergen-Paap rk Wald F-statistic obtained from the first stage of an instrumental variable regression with age at the time of forced displacement as the instrument for educational attainment in Austria vis-à-vis Bosnia.

Table 3.5: Main results

The large F-statistics (>1000) in the first stage of the 2SLS regressions confirm the relevance of age as a strong predictor for having attained host-country education for the sample of among Bosnians aged 13 to 22 when migrating to Austria. Overall, the effect of host country education vis-à-vis similar education attained in the country of origin on the probability of being employed is close to zero on average over the observation period. The coefficient estimated by 2SLS (column 2) implies that host-country education increases the probability of employment by 1.6 percentage points, but the coefficient is not significant at any conventional level. We note that the reason for this small and insignificant effect could be the result of two opposing forces: On the one hand, the higher value of the Austrian degree may encourage employment and labour market participation more broadly as it opens up new opportunities on the labour market. On the other hand, a resulting increase in reservation wages may have a negative effect on employment rates of Austrian vis-à-vis Bosnian degree holders [Reyneri and Fullin, 2011, Zwysen, 2019]. Host-country education has a strong negative effect on both the employment quality outcomes (columns (4) and (6)), confirming the descriptive results of subsection 3.5.1 as well as the coefficients estimated by OLS that are shown in columns (3) and (5) of table 3.5. On average over the observation period, the age at the time of arrival induced host-country education vis-à-vis similar education attained in the country of origin decreases the probability of working in low-skill employment and below educational attainment by 8.5 percentage points and 10.3 percentage points respectively. The confidence intervals are tightly estimated around these coefficients ($p<0.01$).

Appendices 3.C.1, 3.C.2 and 3.C.2 show that these main results are further robust to (i) a Heckman correction for employment participation, (ii) widening and narrowing down the working sample by including and excluding individuals of different age at the time of arrival and (iii) to excluding individuals where the forced displacement may have led to differences in education choices respectively.

Finally, the time spent in the host country is negatively associated with the probability of being employed in low-quality occupations throughout all specifications. We note that the likely reason for the low statistical precision of these estimates is the little arrival-cohort variation these are estimated on when survey year fixed effects are included in the regression; thus, the magnitude of these coefficients should be treated with care. It is further noteworthy that females are less likely to be employed (columns 1 and 2) and are more likely to hold low-quality jobs (columns 3-6). We turn to the characteristics of female employment in more detail in section 3.6.

3.5.3 Quality of employment over time

The estimation of the effect of host-country vis-à-vis origin country education over time is done in two ways. First, equation 3.3 is augmented by a term that interacts the $EducationInAustria_{i,t}$ variable with a continuous "time spent in Austria" variable, measured in years. The latter is then instrumented by a term that interacts the $AgeAtTimeOfArrival_{i,t}$ variable with the same time spent in Austria. The underlying assumption of the approach is that the adjustment path is linear, meaning that any initial discount Bosnian refugees faced on their Bosnian education decreases linearly over time. The results are illustrated in figure 3.7.

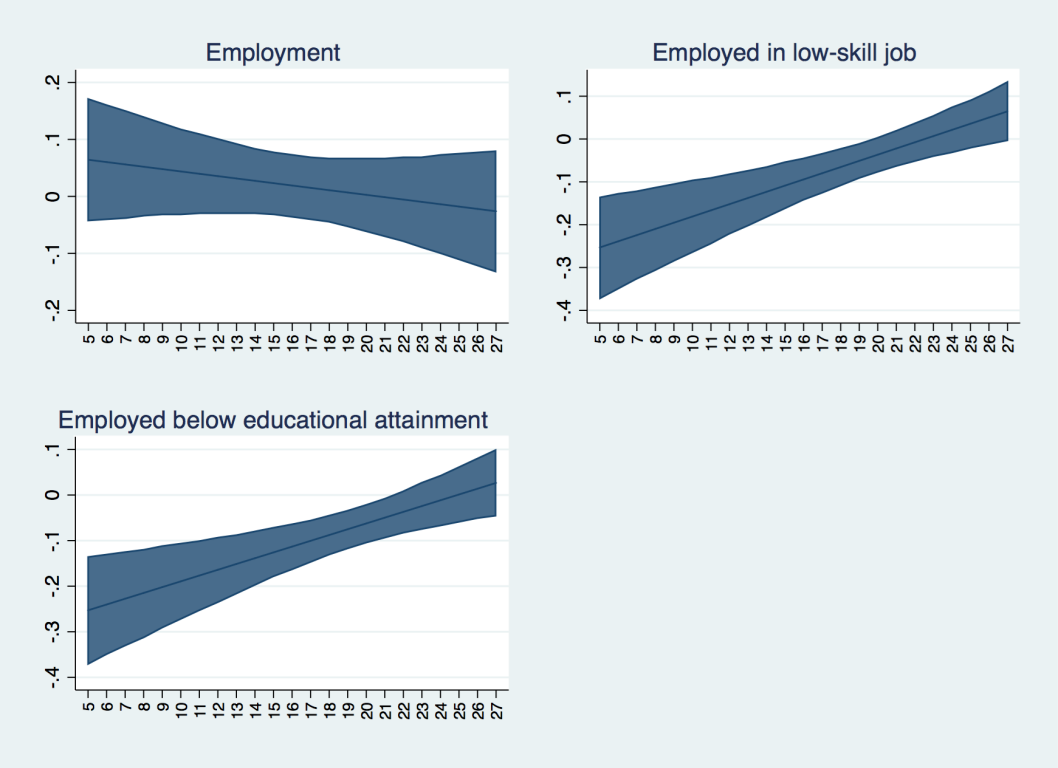


Figure 3.7: The effect of host-country education on employment and employment quality over time of stay

Note: The plots show margins derived from 2SLS regressions for the estimated coefficients on a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia, as well as the interaction term of the indicator with the time spent in Austria (measured in years). The sample consists of individuals aged between 13 and 22 when arriving in Austria during the war in Bosnia that lasted from 1992 to 1995. The observation period is 1998 to 2019. The x-axis shows the time spent in Austria in years. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. Regressions with employment as an outcome are estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. The shaded areas show 90% confidence intervals.

While employment rates among Bosnians educated in Austria do not differ significantly from Bosnians educated in Bosnia at any point of stay, the plots show that the average effects over time mask an adjustment process of employment quality. Both the rates of employment in low-skill occupations (upper right panel) and employment below educational attainment (bottom left panel) among Bosnians with Bosnian degrees start out at a very high level. After five years of stay, the gap to Austrian degree holders stands at 25pp on both indicators. The 90% confidence intervals only start to include zero after about two decades.

To allow for more flexibility in the adjustment over time, a second approach splits the sample by different "time spent in Austria" brackets, similar to figure 3.6. Table 3.6 shows how the results of the preferred 2SLS model.

Table 3.6: The effect of host-country education on employment and employment quality over time of stay - split sample regressions

	Employed			*Work in low-skill job*			*Work below education*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Time spent in Austria	5-12 years	13-20 years	21-27 years	5-12 years	13-20 years	21-27 years	5-12 years	13-20 years	21-27 years
Host-country education	0.086	-0.063	0.077	-0.152**	-0.093**	-0.044*	-0.165**	-0.097**	-0.070**
	(0.059)	(0.050)	(0.065)	(0.066)	(0.043)	(0.026)	(0.068)	(0.041)	(0.033)
Time spent in Austria	0.012	-0.030	-0.012	0.022	-0.020	-0.019	0.037	-0.038*	-0.019
	(0.040)	(0.024)	(0.026)	(0.036)	(0.019)	(0.013)	(0.037)	(0.022)	(0.012)
Female	0.028	-0.112***	-0.017	0.247***	0.235***	0.037***	0.183***	0.139***	0.005
	(0.049)	(0.035)	(0.041)	(0.053)	(0.030)	(0.012)	(0.056)	(0.035)	(0.020)
N	340	573	425	280	453	359	199	327	250
R^2	0.015	0.012	0.009	0.129	0.136	0.029	0.123	0.101	0.014
Estimation method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Time FE	Year	Year	Year	Year	Year	Year	Year	Year	Year
Federal state FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-test	573.537	836.100	331.162	470.033	581.076	240.038	469.522	524.932	196.256

Standard errors clustered at the household level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: The sample consists of individuals aged between 13 and 22 when arriving in Austria during the war in Bosnia that lasted from 1992 to 1995. The observation period is 1998 to 2019. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. Regressions with employment as an outcome are estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. "Host-country education" is a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia. "Time spent in Austria" is a continuous variable measured in years. The first-stage F-test refers to the Kleibergen-Paap rk Wald F-statistic obtained from the first stage of an instrumental variable regression with age at the time of forced displacement as the instrument for educational attainment in Austria vis-à-vis Bosnia.

Columns (1) to (3) again confirm that employment is unaffected by formal host-country education. Columns (4) to (6) and (7) to (9), on the other hand, yield two additional results compared to the pooled sample regressions of table 3.5. First, for Bosnians who stayed in Austria for five to 12 years, holding a comparable Bosnian degree rather than an Austrian degree led to a 15.2pp increase in the probability of low-skill employment for the whole sample of employed Bosnians. For Bosnians educated at an upper secondary level and above, an Austrian degree decreased the likelihood of working below educational attainment by 16.5pp ($p<0.05$). To put these numbers in perspective, consider that between 1998 and 2019, a stable share of close to 6% of native-born Austrians of upper secondary education and above worked below educational attainment. Among Bosnians holding a foreign degree in the working sample, the unconditional share stood at 32% after five to 12 years of stay. Thus, not holding an Austrian degree still explained more than half of the gap in overeducation between Austrians and Bosnian refugees after around a decade in the country. Second, after more time spent in the country (columns 5, 6, 8 and 9), the discount on foreign education decreases until it reaches 4.4pp ($p<0.1$) for the probability of working in a low-skill job and 7pp for the probability to work below educational attainment ($p<0.05$) after more than two decades. These findings are well in line with previous descriptive studies that document a similar convergence in employment quality over the duration of stay among refugees in Europe [Damas de Matos and Liebig, 2014, Zwysen, 2019].

3.5.4 Net hourly wages

The strong positive effect of formal host-country education on employment quality among Bosnian refugees uncovered in subsections 3.5.2 and 3.5.3 is likely to have direct implications for earned income: In 2018, employed Bosnians earned, on average, EUR 1948 in net terms when working in an occupation not classified as low-skill according to equation 3.2. Within low-skill occupations, the average net income stood at EUR 1119. Accounting for differences in hours worked does not change this observation: In 2018, Bosnians in low-skill occupations were paid a net hourly wage of EUR 9.46, compared to EUR 12.46 in non-low-skill occupations.[6]

Data on net monthly income and monthly hours worked are only available from 2011 to 2018 in the Austrian microcensus. Estimates comparable to the other employment indicators can therefore only be produced for the last "time-spent-in-Austria" bracket

[6]Data are unweighted averages calculated based on the 2018 Austrian microcensus.

of table 3.6. These estimates are shown in table 3.7 below for the whole sample of the employed, thus comparable to column (6) of table 3.6 and the sample of Bosnians educated at upper secondary level and above, comparable to column (9) of table 3.6.[7]

	Hourly wage income		*Hourly wage income - no low-educated*	
	(1)	(2)	(3)	(4)
Host-country education	0.087*	0.082	0.125**	0.158**
	(0.045)	(0.058)	(0.053)	(0.068)
Time spent in Austria	0.003	0.003	0.011	0.011
	(0.025)	(0.024)	(0.035)	(0.033)
Female	-0.145***	-0.146***	-0.175***	-0.168***
	(0.040)	(0.038)	(0.052)	(0.050)
N	262	262	175	175
R^2	0.187	0.071	0.247	0.126
Estimation method	OLS	2SLS	OLS	2SLS
Time FE	Year	Year	Year	Year
Federal state FE	Yes	Yes	Yes	Yes
Education category FE	Yes	Yes	Yes	Yes
First-stage F-test		138.312		101.625

Standard errors clustered on the household level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: The sample consists of individuals aged between 13 and 22 when arriving in Austria during the war in Bosnia that lasted from 1992 to 1995. Hourly wage income refers to the logarithm of the hourly wage individuals earned in the month prior to the interview. "Time spent in Austria" is a continuous variable measured in years. The sample consists of individuals who stayed in Austria for between 21 and 27 years and is based on microcensus data from between 2011 and 2018. The first-stage F-test refers to the Kleibergen-Paap rk Wald F-statistic obtained from the first stage of an instrumental variable regression with age at the time of forced displacement as the instrument for educational attainment in Austria vis-à-vis Bosnia.

Table 3.7: Hourly wage income

The results confirm that, even after more than two decades, the age at the time of forced migration induced higher probability to attain formal host-country education led to larger net wages. The estimated coefficient on the "host-country education" indicator in the 2SLS regressions show that net hourly income was 8.2% higher for Austrian degree holders (p<0.2; column 2) among all employed Bosnians and 15.8% higher for Austrian degree holders among employed Bosnian refugees educated at upper secondary level and above (p<0.05; column 4). Interpreting these results in conjunction with the finding of employment quality convergence between Austrian and Bosnian degree holders over time (see table 3.6) further suggests that the income gap between the two groups was even larger in the earlier years of stay.

[7]Missing values in the reported hours worked and monthly income lead to a slightly smaller sample size compared to table 3.6.

3.5.5 The discount on tertiary education

One of the limitations of the working sample of those aged 13 to 22 at the time of forced migration is that it includes a relatively smaller share of tertiary Bosnians with education attained in Bosnia before displacement. Within the sample of Bosnians displaced when aged between 13 and 22, the share of tertiary educated stood at 3.1%, compared to 6.1% among all Bosnians aged between 25 and 64 years that are observed across all microcensus survey years. Thus, while the total number of tertiary educated Bosnians who migrated to Austria was therefore very low, the main working sample still includes relatively fewer tertiary educated. The reasons for the choice of the relatively narrow age bandwidth are related to the exclusion restriction of the suggested IV approach. They are discussed in detail in subsection 3.4.2. A less conservative choice of the bandwidth would therefore partly remedy the issue by including Bosnians that were older at the time of forced migration. Appendix 3.C.2 shows that widening the bandwidth - such that more tertiary educated Bosnians who attained their degrees in Bosnia are included - does not change the main results shown in table 3.5.2; in fact, widening the bandwidth to include Bosnians aged up to 25 at the time of displacement yields point estimates on the employment quality indicators of identical magnitude compared to the main working sample (figure 3.9).

Despite these observations, the high incidence of working below educational attainment among tertiary-educated refugees in Europe justifies a descriptive analysis of the group of tertiary-educated Bosnians in Austria [Dumont et al., 2016]. Table 3.8 shows the results of a linear probability model as formally defined in equation 3.3, but estimated on the pooled sample of tertiary Bosnians aged 25 to 64.[8] Of the 293 total individuals observed, 184 individuals had attained their tertiary degrees in Bosnia and 109 individuals attained their degrees in Austria, after forced displacement.

[8]Unlike in the previous regressions on the smaller subsample of Bosnians aged 13 to 22 at the time of arrival in Austria, controlling for age (and its squared term) is possible in the regressions on the larger sample of individuals aged 25 to 64. However, these controls did not change the point estimates and were therefore left out to ensure comparability across analyses.

	(1) *Employed*	(2) *Low-skill job*	(3) *Work below education*
Host-country education	0.027	-0.037	-0.200**
	(0.068)	(0.027)	(0.084)
Time spent in Austria	0.036	-0.025	-0.022
	(0.040)	(0.028)	(0.048)
Female	-0.052	0.060**	0.049
	(0.059)	(0.026)	(0.068)
N	293	225	225
R^2	0.090	0.209	0.260
Estimation method	OLS	OLS	OLS
Time FE	Year	Year	Year
Federal state FE	Yes	Yes	Yes

Standard errors clustered at the household level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: The sample consists of individuals aged between 25 and 64 that are educated at a high level (ISCED-97 level above 4). The observation period is 1998 to 2019. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. The variable "work in low-skill job" takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The variable "work below education" takes the value 1 if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. "Host-country education" is a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia. "Time spent in Austria" is a continuous variable measured in years.

Table 3.8: Tertiary educated

Despite the relatively small sample of tertiary-educated Bosnians, the results uncover some noteworthy findings. First, the probability of employment is again not associated with whether education was attained over the pooled observation period. Second, while tertiary-educated Bosnians with degrees from Bosnia are indeed 20pp more likely to work below their educational attainment than Bosnians completing tertiary education in Austria (column 3), their downgrading appears to happen mostly into medium-skill, rather than low-skill jobs (column 2).

The results of table 3.8 thus confirm previous analyses that document a large discount refugees face on their tertiary education [Damas de Matos and Liebig, 2014, Dumont et al., 2016].

3.6 Cultural legacy: Differentiating male and female refugees and the role of female employment in former Yugoslavia

One of the striking features throughout the analyses in section 3.5 is the large magnitude of the negative coefficient on the female migrant indicator in regressions related to the quality

of employment (see for example table 3.5). This difference in labour market integration outcomes and trajectories between male and female migrants is often observed and has led scholars to routinely analyse these groups separately [Dustmann, 1997, 1999].

In the context of forced migration from Bosnia to Austria and the research question of refugee labour market integration, this is particularly interesting for a number of reasons. First, in 1991, the year before the start of the war, estimates from the International Labour Organisation (ILO) indicate that female labour market participation of people aged 15 and above stood at 41% (males: 63%) in Bosnia. The female-to-male ratio of labour market participation thus closely resembled that of Austria in the same year (female: 44%; male: 70%). The relatively high labour market participation rate among women was most likely a relic from socialist times in the former Yugoslavia, where workplaces often provided housing, child care and health care [Pascall and Manning, 2000]. A second feature of female employment is stressed in historical research on gendered work in former Yugoslavia: Women were predominantly employed in low-skill jobs and rarely held high positions in factories or politics [Pascall and Manning, 2000, Darville and Reeves, 1992].

In the following, we therefore analyse whether these pre-migration characteristics of female employment are observed on the Austrian labour market and if the attainment of host-country education could alter outcomes vis-à-vis those who entered the Austrian labour market immediately upon arrival and attained their education in Bosnia.

Table 3.9 shows the regression results for female migrants and male migrants respectively.

	Females			*Males*		
	(1) Employed	(2) Low-skill job	(3) Below educ.	(4) Employed	(5) Low-skill job	(6) Below educ.
Host-country education	0.003	-0.130***	-0.159***	0.037	-0.024	-0.042**
	(0.047)	(0.048)	(0.050)	(0.046)	(0.019)	(0.020)
Time spent in Austria	-0.014	-0.015	-0.018	0.006	0.004	-0.003
	(0.021)	(0.019)	(0.022)	(0.020)	(0.005)	(0.007)
N	775	612	387	579	491	396
R^2	0.001	0.018	0.063	0.005	0.006	0.016
Estimation method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Time FE	Year	Year	Year	Year	Year	Year
Federal state FE	Yes	Yes	Yes	Yes	Yes	Yes
Education category FE	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-test	974.654	713.255	631.615	695.869	512.956	454.371

Standard errors clustered at the household level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: The sample consists of individuals aged between 13 and 22 when arriving in Austria during the war in Bosnia that lasted from 1992 to 1995. The observation period is 1998 to 2019. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. Regressions with employment as an outcome are estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. "Host-country education" is a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia. "Time spent in Austria" is a continuous variable measured in years. The first-stage F-test refers to the Kleibergen-Paap rk Wald F-statistic obtained from the first stage of an instrumental variable regression with age at the time of forced displacement as the instrument for educational attainment in Austria vis-à-vis Bosnia.

Table 3.9: Main results split by males and females

It is striking that the total observed effect of host-country education is particularly driven by female Bosnians. Just like for males, the effect of host-country education vis-à-vis education attained in Bosnia is small and not clearly distinguishable from a zero effect. However, the magnitude of the effect of Austrian vis-à-vis Bosnian education on the employment quality indicators is striking: For Bosnian women, host-country education decreases the probability of working in low-skill employment and below educational attainment by 13.0pp and 15.9pp respectively on average over the 22 year observation period (columns 2 and 3). A closer look at the data reveals that in any given year of the observation period, Bosnian women who entered Austria when aged between 18 and 22 were predominantly employed in the domestic work and cleaning sectors, while their female counterparts aged 13 to 17 upon arrival worked in a wide range of professions.

On the other hand, host-country education appears to have less of an effect on the employment quality of male migrants on average (columns 5 and 6). The sign of the estimated coefficients is still negative but the magnitude of these coefficients is smaller and conventional confidence intervals now include zero for the "low-skill job" indicator (column 5). Host-country education still decreases the probability of working below educational attainment by 4.2pp for males (p<0.05), but the magnitude of the coefficient is less than

30% of the effect in females.

Figure 3.8 shows that these differences only emerge in the later years of refugees' stay in Austria.

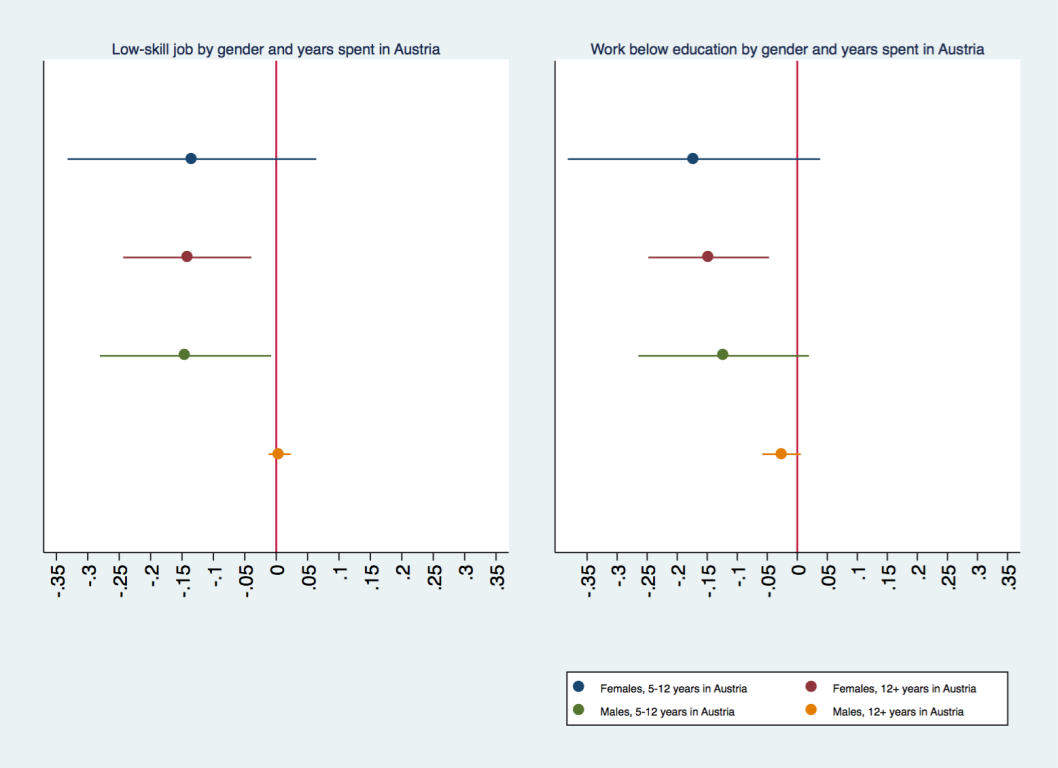


Figure 3.8: Employment quality - females and males over time of stay

Note: The figure plots the coefficients from the same regressions as in table 3.9 but splitting the sample into females and males who had spent 5 to 12 years in Austria and females and males who spent more than 12 years in Austria.

While after five to 12 years of stay in Austria, both male and female Bosnian refugees faced a large discount on their education, men managed to close the gap in the later years of stay while females did not. The cultural legacy of working in low-skill occupations among women in former Yugoslavia may have played a role for the stickiness of low-quality employment: While both men and women faced an initial discount on their education in Bosnia, likely due to a mix of quality differences in education between Bosnia and Austria and the related lower signalling value of Bosnian degrees, women faced the additional difference in female work culture between the countries. While agnostic on the precise mechanisms at play, these results thus lend strong support to the hypothesis that host-country education can help overcome cultural barriers to the labour market. Additionally, the results could reflect that Bosnian women were more strongly affected by the institutional setting that incentivised them to enter into employment shortly upon arrival.

Qualitative research on the situation of Bosnian women in Austria during the Bosnian war years conducted by Franz [2005] emphasises the important role of Bosnian women in securing "their families' residence rights as well as such social benefits as health insurance" (p.55) when Bosnians first came to settle in Austria. Thus, the institutional structure in Austria that discouraged humanitarian migrants from attaining formal education beyond a certain age may well have exacerbated existing cultural effects.

3.7 Concluding discussion

In this paper, I contribute to the question of how labour market outcomes of humanitarian migrants are shaped by the institutional setting they migrate into. Using the example of Bosnian refugees in Austria, I show that, vis-à-vis similar levels of education attained from the country of origin, a degree from the host country dramatically improves the labour market position of humanitarian migrants.

The causal results obtained from instrumental variable regressions recover local average treatment effects only valid among the young humanitarian migrant population. However, young humanitarian migrants from less developed regions seeking asylum in EU member states were not only a highly relevant group in the past but also feature prominently in current inflows of asylum seekers. 81% of the 5.7 million migrants who lodged an asylum application in an EU member state between 2008 and 2018 were aged below 35, originating predominantly from Syria, Afghanistan and Iraq. The gap between the quality of education differs widely between EU destination countries and these more recent countries of origin [Hanushek and Woessmann, 2011, Bonfanti and Xenogiani, 2014]. For these inflows, the estimates obtained in this paper are thus likely to present a lower bound. Supporting asylum seekers to attain host-country education would likely have larger payoffs to migrants when the origin-destination distance in human capital is large; on the other hand, the public investment required is a similarly positive function of this distance.

It is important to note that the analyses in the paper do not require humanitarian migrants to upgrade their education. The study makes the simple point that, if humanitarian migrants were to be integrated into the host-country education system until they achieve the same qualification they had previously attained in their country of origin, this would dramatically improve their position on the labour market. The estimates should thus not be interpreted as returns to schooling estimates as in, for example, Oreopoulos

[2006], but as explicitly comparing the value of education in the country of origin to the destination country for humanitarian migrants. Thus, they combine the theoretical effects of quality differences in education, signalling, screening and potential peer effects.

A natural follow-up question is to what extent it is realistic to integrate humanitarian migrants into the host-country education system swiftly upon arrival. An efficient policy approach would allocate significant resources into thorough assessments of existing skills. Austrian authorities did not invest into procedures to assess and recognise degrees from Bosnia when refugees started arriving during the Bosnian war, leaving many new-arrivals with few options on the labour market [Tretter, 2000]. An acknowledgement of foreign degrees or an assessment of shortcomings vis-à-vis an equivalent degree in the host country could be followed by training in host-country specific language skills. Humanitarian migrants would then be able to enter the host-country education system at an appropriate level.

In this context, the external validity of the results found in this study require a careful reflection: Depending on the design of policy interventions that aim to increase host-country education among refugees, the treatment effects may differ depending on policy parameters that determine the depth of integration into the host-country education system. For example, in Austria, no parallel education system existed that kept humanitarian migrants separate from the native population [Tretter, 2000]. If, instead, migrants are taught in classes separate from the host-country population, this could potentially diminish the positive effect on their future performance on the labour market. The literature on peer effects in schools shows that these are generally complex across different student subgroups [Lavy et al., 2012]. However, network formation with host-country students may be relatively more important for humanitarian migrants than for the general population. Further research is required on the mechanism that guides the impact of host-country education on future labour market performance through exposure to native peers.

A key finding of this study is that host-country education benefitted female Bosnian refugees significantly more regarding their labour market position than their male counterparts. Attaining education in the host-country drastically remedies the education-employment mismatch for women. The cultural background of Bosnians, where women traditionally held low-level positions on the labour market may thus well have translated into worse employment outcomes in Austria, where the institutional setting discouraged host-country education. The finding suggests that the education system can play a decisive role in overcoming cultural barriers in the labour market. This is important going

forward as the gender dimension of labour market integration is likely to play an even more important role for more recent inflows of humanitarian migrant in Europe. For example, less than 15% of Syrian women were economically active prior to the breakout of the civil war in Syria that caused more than a million Syrians to flee to Europe [Barslund et al., 2017b]. Policy measures specifically targeted towards women are thus inevitable and host-country education could play a crucial role in the efficient labour market integration of female humanitarian migrants.

The findings of this study further raise the question why even young humanitarian migrants, who are (made) aware that their education is discounted upon arrival, may still decide not to invest into formal host-country human capital. The answer is twofold and lies in the incentive structures created by institutions in the European context. First, humanitarian migrants face significant uncertainty regarding the duration of stay in EU host-countries which has been shown to be an important determinant of migrants' human capital investment decisions [Cortes, 2004, Dustmann, 1997, 1999]. If migrants expect to stay in the host-country with high certainty, they undertake costly investments into human capital as they anticipate that short-term foregone income will be more than compensated for by higher future earning potential [Duleep and Regets, 1999, Cortes, 2004]. In the EU, a large share of humanitarian migrants faces uncertainty regarding the length of their stay in EU member states. Upon arrival, asylum seekers go through asylum procedures of unknown outcome where even the first instance lasts between six and twelve months on average [Commission, 2016a].[9] Since only 50 percent of asylum seekers ultimately gain humanitarian protection, it is rational to discount the probability of an extended stay. After these asylum procedures, the share of asylum seekers receiving full Geneva Convention refugee status, the status that grants humanitarian migrants the most rights, is small in the EU and stood at less than 25 percent of all asylum applicants over the past decade.[10] Thus, even for those 50 percent of asylum seekers who ultimately obtain protection, this does not necessarily mean certainty: A range of different protection regimes exist across the EU and even in the best possible outcome when asylum seekers gain full Geneva convention refugee status, the length of stay is far from certain.[11]

[9] If an asylum decision is rejected, asylum seekers have the chance to appeal this decision. Around 13 percent of all asylum applications were granted after the asylum seeker had appealed the initial rejection between 2008 and 2017 (Eurostat data).

[10] Calculations based on Eurostat data, 2008-2017.

[11] Residence permits granted to recognised Geneva Convention refugees are initially limited to three to five years in the majority of European countries [Commission, 2016a]. Very large cross-country differences in the management of humanitarian migrants of the same protection status across EU member states further add to the uncertainty regarding the probability of stay. For example, when Bosnian refugees entered various European countries following the war in Bosnia in the early 1990s, they initially received

Second, if there is a policy preference towards an economic integration and insufficient infrastructure to support humanitarian migrants to integrate into education systems, this presents a relative increase in the cost of human capital investment compared to taking up low-skill labour.[12]

Finally, this paper focuses on the economic aspects of integrating humanitarian migrants into host-country education systems. Policymaking may benefit from a deeper understanding of the long-term societal costs a potential systematic clustering of specific migrant groups in predominantly low-skill, low-income occupations incurs. For example, studies have analysed the relation between immigrants' access to the labour market and crime rates, finding a strong negative link between the two [Freedman et al., 2018, Couttenier et al., 2019]. Future research could specifically look into the effect of the clustering of migrants in low-quality employment on measures of social cohesion in the host country.

temporary protection in all host countries. As soon as the war ended in 1995, Germany repatriated almost all refugees. On the other end of the spectrum, Sweden granted refugees permanent residency soon after arrival. Other main recipient countries such as Austria, Denmark or the Netherlands fell in between [Barslund et al., 2017a].

[12]In Europe, the prioritisation of integrating humanitarian new arrivals into the labour market over an integration into the host-country educational system manifests itself in several practical dimensions. In its 2019 evaluation of the Recast QD, the European Commission concludes that 20 out of 29 EU member states grant refugees access to adult education under the same conditions as legally residing adult third-country nationals [Commission, 2016b]. However, unlike employment assistance, only a "limited number of Member States have provided beneficiaries of international protection with additional support to facilitate access to the national education system" [Commission, 2016b, p.213]. Neither the European Statistical Office (Eurostat) nor any EU member state currently collects data on the number of beneficiaries of humanitarian protection accessing host-country education, with the exception of Slovenia where the number of beneficiaries in the country is small and the total number of adults and minors accessing education stood at 62 in 2015. Reports on more specific aspects of integration into the education system reach similar conclusions. For example, a recent review of asylum seekers' access to (higher) education by the European Commission concludes that "the majority of [EU] countries have no specific policy approach to integrate asylum seekers and refugees into higher education" [Crosier and Kocanova, 2019, p.24].

Appendix

3.A Facilitators of early employment of Bosnian refugees

Two mediating channels may have had a positive impact on the integration of Bosnian refugees into Austrian labour markets. First, due to large guest worker movements from Bosnia in the preceding decades, Bosnians had extensive co-ethnic networks to draw on all across Austria; even during the peak time of humanitarian inflows into Austria, only about 50 percent of Bosnians relied on publicly organised accommodation [Bendl, 2014]. In total, about 198,000 foreigners born in the former Yugoslavia resided in Austria before the start of the war in 1991 [Fassmann and Reeger, 2008]. The shares in table 3.10 below are calculated based on Austrian microcensus data from 1995 and therefore capture only those Bosnians still present in the country in 1995 who had migrated to Austria before 1992.

Federal state (NUTS-2)	Population share (1991)	Share Bosnians (1991)	Share Bosnian refugees
Burgenland	0.04	0.02	0.02
Kärnten	0.07	0.06	0.07
Niederösterreich	0.19	0.1	0.09
Oberösterreich	0.17	0.23	0.25
Salzburg	0.06	0.12	0.1
Steiermark	0.15	0.1	0.13
Tirol	0.08	0.07	0.08
Vorarlberg	0.04	0.04	0.01
Wien	0.2	0.27	0.25

Notes: Table based on the total sampled population by federal state and unweighted counts of Bosnians who had migrated to Austria in 1991 or before as reported in the 1995 Austrian microcensus. Columns (1), (2) and (3) show the share of the Austrian population residing in each federal state, the share of Bosnians who arrived in 1991 or before in total Bosnians who arrived in 1991 or before residing in each federal state and the share of Bosnian refugees in total Bosnian refugees residing in each federal state respectively.

Table 3.10: Bosnians in Austria before the start of the Bosnian war

The existence of co-ethnic networks in the destination country have been linked to higher earnings among refugees in previous research. In particular informal channels facilitate better job market matches [Damm, 2009, Edin et al., 2003]. On the other hand,

Battisti et al. [2016] show that the existence of co-ethnic networks not only leads to a faster integration of refugees into the labour market, but is also linked to lower host-country human capital investments. It remains unclear if co-ethnic networks are structural parameters in the human capital investment decision or should rather be considered a facilitator of an underlying preference for employment. In this context, it should further be noted that the large co-ethnic networks Bosnians could draw on in Austria may have led to a negative selection of refugees from Bosnia along their education levels [Beine et al., 2011].

Second, the labour market conditions in Austria were extremely favourable throughout the 1990s with very small regional and annual deviation from the national mean unemployment rate of around 4 percent. Official Austrian statistics did not report regional unemployment rates but the total number of unemployed by federal state. Table 3.11 shows calculated unemployed-to-population ratios for each federal state. Thus, unemployment was largely low and homogenous across the country. These favourable labour market conditions likely benefitted Bosnians in particular in the beginning when barriers to the labour market were still in place.

Federal state (NUTS-2)	Unemployment-to-population 1992-1995
Burgenland	2%
Kärnten	3%
Niederösterreich	2%
Oberösterreich	2%
Salzburg	2%
Steiermark	3%
Tirol	2%
Vorarlberg	2%
Wien	4%

Notes: Calculations based on the number of unemployed and total federal state population reported by Statistics Austria. Statistics averaged over the 1992 to 1995 period.

Table 3.11: Unemployment-to-population ratios in Austrian federal states between 1992 and 1995

3.B Labour market access of Bosnian refugees in Austria during and after the Bosnian war

A wider access to the labour market was granted with some initial limitations as Bosnian refugees were initially placed last in a priority system that favoured native-born Austrians, EU migrants and other third-country nationals. Table 3.12 provides a comprehensive

overview of the timeline of labour market access that was gradually granted to Bosnian refugees in Austria. The initial restrictions led to informal employment in the early years [Franz, 2003]. However, formal employment rates then picked up quickly [Barslund et al., 2017a].

1992	1993	1994	1995	1996-1998
July: First access via a special contingent but only temporary employment of up to three months with municipalities and registered charities October: The above time limit gets gradually extended but still only municipalities and charities legally allowed to hire Bosnians	July: Full access to the labour market but within a priority system that put Bosnians last after Austrians and other migrants	August: Still last in the priority system but Austrian employment services start to support Bosnians in finding work	First work permits issued; initially only 2000 but steadily increasing	All Bosnians obtain work permits
Maximum quota of foreign workers not to exceed 8-9% of the total work force				

Notes: Table based on Tretter [2000].

Table 3.12: Labour market access of Bosnian refugees in Austria over time

An important institutional feature that incentivised Bosnians to integrate into the formal labour market quickly despite some early barriers was the fact that their residence status could be converted into a more permanent status through employment if earnings where sufficient to sustain themselves and their families. Around 30,000 to 40,000 Bosnian refugees managed to get a permanent residence status through employment under this so-called Aliens Act [Tretter, 2000].

3.C Further robustness tests

This section presents further robustness tests. Subsection 3.C.1 shows the results when selection into employment is accounted for. Subsection 3.C.2 shows the results for different definitions of the working sample. Subsection 3.C.3 analyses the effect of the forced displacement experience on the level of educational attainment in Austria. Finally, sub-

section 3.C.4 discusses the validity of the IV approach in light of the importance of the external versus the internal margin of schooling in Austria.

3.C.1 Heckman correction for employment participation

We first use the entire sample of Bosnians, including those not in employment, to estimate a selection equation of the form

$$Probit: Pr(Employed = 1|x_i)_{i,t} = \Phi(\Pi_0 + \Pi_1 EducationInAustria_{i,t} + \delta Z_{i,t} + \xi_i), \quad (3.5)$$

where x_i is a list of regressors and Φ is the cumulative distribution function of the standard normal distribution. We further note that the matrix of control variables, $Z_{i,t}$, differs slightly from $IndividualCharacteristics_{i,t}$ in equation 3.3 since we are now explicitly modelling participation in employment instead of employment quality. The general decision to obtain employment may, in addition to the individual level characteristics, also depend on household characteristics such as the relationship status, the number of persons in the household or the number of children in the household. These variables are used to identify 3.5. Due to some missing values in these variables, the sample size is reduced slightly.

Obtained from equation 3.5, the estimated inverse Mills ratio (IMR) denoted by λ is a standard hazard function computed for each individual i by

$$\hat{\lambda}_i = \frac{-\phi(x_i)}{\Phi(x_i}, \quad (3.6)$$

where ϕ is the probability density function. $\hat{\lambda}_i$ is then included in equations 3.3 and 3.4 as a regressor. Since the IMR is estimated and not known, standard errors in the regressions are cluster bootstrapped, based on 1000 bootstrap iterations.

The results of the regressions that include the Heckman corrections are shown in table 3.13.

	Employed		*Work in low-skill job*		*Work below education*	
	(1)	(2)	(3)	(4)	(5)	(6)
Host-country education	0.045*	0.016	-0.051**	-0.087***	-0.069***	-0.090***
	(0.025)	(0.033)	(0.024)	(0.031)	(0.025)	(0.031)
Time spent in Austria	-0.005	-0.005	-0.007	-0.007	-0.010	-0.011
	(0.015)	(0.015)	(0.011)	(0.011)	(0.011)	(0.011)
Female	-0.041*	-0.045*	0.147***	0.147***	0.076***	0.077***
	(0.024)	(0.024)	(0.020)	(0.020)	(0.022)	(0.022)
Lambda			0.326***	0.255*	0.512***	0.463***
			(0.119)	(0.130)	(0.142)	(0.153)
N	1354	1354	1089	1089	775	775
R^2	0.048	0.005	0.246	0.087	0.263	0.097
Estimation method	OLS	2SLS	OLS	2SLS	OLS	2SLS
Time FE	Year	Year	Year	Year	Year	Year
Federal state FE	Yes	Yes	Yes	Yes	Yes	Yes
Education category FE	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-test		1743.729		1729.983		1460.857

Standard errors cluster bootstrapped at the household level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: The sample consists of individuals aged between 13 and 22 when arriving in Austria during the war in Bosnia that lasted from 1992 to 1995. The observation period is 1998 to 2019. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. Regressions with employment as an outcome are estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. "Host-country education" is a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia. "Time spent in Austria" is a continuous variable measured in years. "Lambda" is the inverse Mills ratio obtained from a probit regression formally described in 3.5 and corrects for employment participation. The first-stage F-test refers to the Kleibergen-Paap rk Wald F-statistic obtained from the first stage of an instrumental variable regression with age at the time of forced displacement as the instrument for educational attainment in Austria vis-à-vis Bosnia.

Table 3.13: Baseline results with Heckman correction for employment participation

3.C.2 Adjusting the age bandwidth

The sample selection of 12 to 23 year olds is an arbitrary choice that balances the validity of the instrument against the sample size. Figure 3.9 shows that the main results are robust to different choices of the bandwidth.

Figure 3.9: The effect of host-country education on employment and employment quality for different bandwidths

Note: The plots show margins derived from 2SLS regressions as formally described in equations 3.3 and 3.4 for the estimated coefficients on a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia. The bandwidths chosen correspond to the age at the time of arrival in Austria. The observation period is 1998 to 2019. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. Regressions with employment as an outcome are estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. The horizontal bars show 95% confidence intervals.

3.C.3 Accounting for potential differences in educational attainment by age

Figure 3.10 shows the conditional shares of low-educated, medium-educated and highly-educated Bosnians in Austria by the precise age at forced migration. The green line shows the sample mean for each education group among Bosnians aged 10 to 30 years at the time of migration.

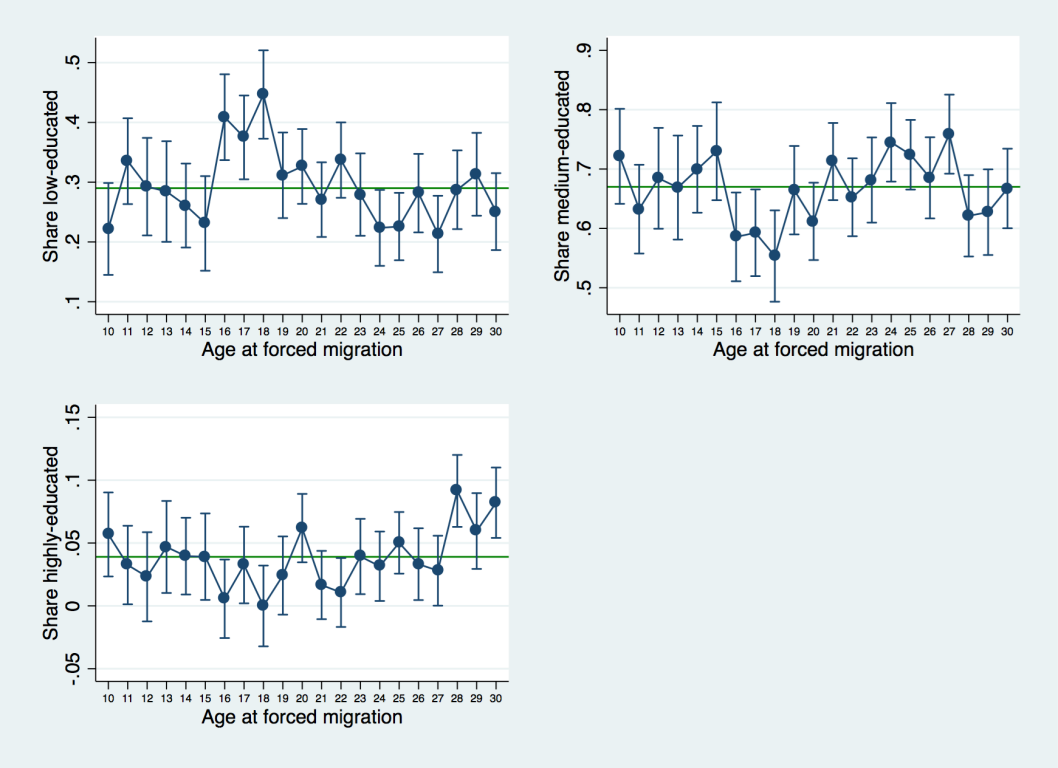


Figure 3.10: Educational attainment by age

Note: The sample for the shown margins consists of all individuals aged between 10 and 30. Margins are derived from an OLS regression of a binary indicator that takes the value one if an individual fell into the respective education category (and zero otherwise) on a fixed effects term capturing age at the time of forced displacement. The estimates are conditioned on a dummy variable for females, a continuous indicator for time spent in Austria (in years), survey year and federal state fixed effects. The green lines show the sample mean for each education group. The vertical bars show 95% confidence intervals.

The figure shows that Bosnians aged 16, 17 and 18 at the time of forced displacement were significantly more likely to be educated at a low level (upper left panel) than at a medium level (upper right panel), compared to the sample average of Bosnians aged 10 to 30 years at the time of migration. Excluding individuals aged 16, 17 and 18 at the time of forced displacement yields the results shown in table 3.14.

	Employed		*Work in low-skill job*		*Work below education*	
	(1)	(2)	(3)	(4)	(5)	(6)
Host-country education	0.042	0.023	-0.066***	-0.062**	-0.087***	-0.091***
	(0.028)	(0.033)	(0.024)	(0.028)	(0.025)	(0.030)
Time spent in Austria	0.002	0.002	0.001	0.001	-0.009	-0.009
	(0.018)	(0.018)	(0.013)	(0.012)	(0.013)	(0.013)
Female	-0.045	-0.048*	0.202***	0.202***	0.138***	0.137***
	(0.028)	(0.028)	(0.024)	(0.023)	(0.026)	(0.026)
N	996	996	807	807	590	590
R^2	0.054	0.006	0.255	0.102	0.252	0.077
Estimation method	OLS	2SLS	OLS	2SLS	OLS	2SLS
Time FE	Year	Year	Year	Year	Year	Year
Federal state FE	Yes	Yes	Yes	Yes	Yes	Yes
Education category FE	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-test		2235.278		1524.061		1249.213

Standard errors clustered at the household level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: The sample consists of individuals aged between 13 and 22 when arriving in Austria during the war in Bosnia that lasted from 1992 to 1995 but excludes those aged 16, 17 and 18. The observation period is 1998 to 2019. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. Regressions with employment as an outcome are estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. "Host-country education" is a dummy variable that takes the value 1 if the highest education was attained in Austria and the value 0 if it was attained in Bosnia. "Time spent in Austria" is a continuous variable measured in years. The first-stage F-test refers to the Kleibergen-Paap rk Wald F-statistic obtained from the first stage of an instrumental variable regression with age at the time of forced displacement as the instrument for educational attainment in Austria vis-à-vis Bosnia.

Table 3.14: Main results excluding 16 to 18 year olds

3.C.4 External versus internal margin of schooling

The instrumental variable approach described in section 3.4.2 makes the implicit assumption that the age at the time of forced displacement is a valid instrument for having received attained a degree from Austria vis-à-vis Bosnia (the external margin). This implies that the years of schooling in Austria (the internal margin when total schooling is held constant), does not predict employment quality in the particular setting at hand. The years of schooling received is not directly reported in the Austrian microcensus such that these regressions cannot be run using a 2SLS regression model. However, to test the assumption, we can restrict the sample to humanitarian migrants who attained formal education in Austria and specify a model of the form

$$Y_{i,t,e} = \beta_t + \kappa_e + \gamma AgeWhenEnteringAustria_{i,t} + \zeta IndividualCharacteristics_{i,t} + \epsilon_{i,t}, \tag{3.7}$$

where γ, the coefficient estimated on the continuous $AgeWhenEnteringAustria_{i,t}$ variable is the coefficient of interest as it may potentially predict the number of years of schooling individuals received. Model 3.7 includes the same controls as model 3.3. It is estimated by OLS.

Table 3.15 shows the results for the working sample, aged 13-22 at forced migration, and a sample of 10 to 14 year olds at the time of forced migration, where the linear $AgeWhenEnteringAustria_{i,t}$ variable serves as a more precise proxy for the years of schooling attained in Austria vis-à-vis Bosnia.

	Aged 13-22 at forced migration			*Aged 10-14 at forced migration*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Employed	Low-skill job	Below educ.	Employed	Low-skill job	Below educ.
Age when entering Austria	0.008	-0.001	0.006	0.015	0.004	0.005
	(0.006)	(0.005)	(0.006)	(0.011)	(0.009)	(0.009)
Time spent in Austria	-0.018	0.005	-0.001	0.002	-0.006	0.005
	(0.019)	(0.011)	(0.014)	(0.024)	(0.014)	(0.010)
Female	-0.048	0.100***	0.045*	-0.077**	0.062**	0.066***
	(0.032)	(0.023)	(0.025)	(0.035)	(0.025)	(0.024)
N	630	535	476	582	472	397
R^2	0.075	0.197	0.257	0.106	0.145	0.134
Estimation method	OLS	OLS	OLS	OLS	OLS	OLS
Time FE	Year	Year	Year	Year	Year	Year
Federal state FE	Yes	Yes	Yes	Yes	Yes	Yes
Education category FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors clustered at the household level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: The sample consists of individuals aged between 13 and 22 (columns 1-3) or between 10 and 14 (columns 4-6) when arriving in Austria during the war in Bosnia that lasted from 1992 to 1995. The observation period is 1998 to 2019. Employed is a dummy variable taking the value 1 if an individual is employed, and 0 otherwise. Regressions with employment as an outcome are estimated on the whole sample of Bosnians. The outcome "work in low-skill job" is estimated on the sample of all employed individuals. The variable takes the value 1 if the employed individual works in an occupation that primarily employs low-educated workers as formally defined in equation 3.2. The outcome "work below education" is estimated on the sample of all employed individuals with at least a medium level of education. The variable takes the value one if the individual works in an occupation that is primarily carried out by workers of lower educational attainment as formally defined in equation 3.1. "Age when entering Austria" is a continuous variable. "Time spent in Austria" is a continuous variable measured in years.

Table 3.15: Years of schooling

No significant association between the age when entering Austria and employment or employment quality indicators is detected. This holds for the working sample (columns 1-3) and the sample aged 10 to 14 (columns 4-6).

Chapter 4

Global food prices and migration in Sub-Sahara Africa

4.1 Introduction

Variability in global food prices has increased significantly over the past two decades: The annual standard deviation around the decade mean of the FAO food price index tripled in the 2000s compared to the 1990s and remains high until today. This phenomenon became most visible during the global food crisis of 2007/08, when international prices of most commodities, including staple grains, increased to an all time high [von Braun, 2008, Minot, 2010]. Over the next decades, the trend is likely to continue. Climate variability, responsible for approximately one third of global crop yield variation between 1979 and 2008, is expected to increase further [Ray et al., 2015]. The resulting variation in global agricultural output will significantly affect international food prices.

In this paper, we study how these increasingly volatile global crop prices affect the household decision to send a member as a migrant in Sub-Sahara Africa. We start out by positing that the main mechanism through which fluctuations in locally-relevant global prices may affect these decisions is the direct effect of crop prices on agricultural households' real income. Several recent studies have provided strong evidence that global crop prices pass through to local agricultural producers and consumers in developing countries, leading to exogenous shocks to households' real income [Alem and Söderbom, 2012, Ivanic et al., 2012, Verpoorten et al., 2013, Hallegatte and Rozenberg, 2017, McGuirk and Burke, 2020]. The aggregate impact of short-term income shocks on the migration decision is then determined through an interplay of two opposing forces: Households' ability to bear

the up-front costs of migration on the one hand, and the opportunity costs of migration that increase with rising income levels on the other. Exogenous variation in household income has therefore been shown to be heterogenous across the household wealth distribution [Clemens, 2014, Cattaneo and Peri, 2016, Hirvonen, 2016, Bazzi, 2017, Dao et al., 2018]: For the credit-constrained poorest households, a positive (negative) income shock increases (decreases) the propensity to migrate as liquidity constraints become less (more) binding. For relatively wealthier households that are not constrained in their ability to send migrants, a rise (fall) in income makes migration less (more) attractive. The main focus of this study is therefore on the heterogeneous effects of locally-relevant global crop prices on the migration decision along the household wealth distribution.

To test the hypothesis, we first build a unique household panel dataset for Burkina Faso, Kenya, Nigeria, Uganda and Senegal, five countries in Sub-Sahara Africa characterised by low average incomes and a high average share of households depending on agricultural production. Thus, changes in agricultural prices in these countries significantly affect real income of crop-producing households. We use retrospectively reported data on internal and international migration to construct an annual measure of out-migration on the household level for the years 2000 to 2008, differentiating between agricultural and non-agricultural households. We then regress this measure on a producer price index (PPI) for agricultural products we construct at the district-year level. Following McGuirk and Burke [2020], the PPI combines high-resolution, time-invariant spatial data on crop-specific agricultural land cover from 2000 with annual international commodity price data over the subsequent years. To benchmark our results against local crop growing conditions and to add precision to our estimates, we complement our data with information on daily temperature and monthly precipitation averages during the growing season. Finally, we further explore a possible mediating role of violent conflict triggered by fluctuations in food prices, a link that has been documented in recent studies [Bellemare, 2015, McGuirk and Burke, 2020, De Winne and Peersman, 2019].

Our results show that fluctuations in locally-relevant global crop prices indeed impact the probability of agricultural households to send a household member as a migrant. The effect is positive on aggregate, implying the existence of a household budget constraint that becomes binding as income from agricultural production falls. On average, 7.6% of agricultural households sent out a household member as a migrant in a given year over our observation period. The estimated coefficients obtained from our preferred specification show that a one standard deviation increase in the PPI over its long-run mean increases the probability of observing household out-migration by 1.3 percentage points. Benchmarking

the magnitude of these effects against those of local weather conditions during the crop-growing season shows that a one standard deviation increase in locally-relevant global crop prices increases household out-migration by 37% of the net effect of a comparable local weather shock. This main result is further corroborated when analysing the effect of the PPI along the household wealth distribution: For the poorest agricultural households, a rise in producer prices strongly raises the probability of household out-migration in a given year; for the relatively wealthier households, there is no link between the PPI and the probability of a household member migrating. When turning to destination choices, we find that, unlike positive weather shocks, which mostly facilitate internal rural-urban migration, positive income shocks through rising producer prices only increase migration to neighbouring African countries. We posit that this finding is likely due to the simultaneous decrease in real income in nearby urban areas caused by rises in crop prices.

Four reasons allow for a plausibly causal interpretation of the coefficients on global food prices in this particular setting: First, our empirical analysis incorporates household fixed effects such that our coefficients of interest capture the effect of global price deviations from their location specific long-term mean over time, thus allowing us to compare a given household under different global price regimes. Second, from 1989 to 2013, the entire continent of Africa accounted for only 5.9% of global cereal production, minimising its effect on global prices [McGuirk and Burke, 2020]. Our analyses further focusses on the observation period from 2000 to 2008, a time span covering the global food price crisis of 2007/08, when global food prices rose sharply for reasons entirely exogenous to agricultural activity in Sub-Sahara African [Berazneva and Lee, 2013, Demeke et al., 2009, Dorward, 2012]. To further support the argument that global crop prices move independently of production in Sub-Sahara Africa, we show in all our analyses that the effect of global prices on our outcomes are not sensitive to conditioning estimates on the local quantity produced, proxied by local weather conditions. Third, our locally-relevant global price measure holds the crop-specific agricultural land cover constant in the first year of our observation period. The approach closely resembles a Bartik-style shift-share instrument pioneered by Bartik [1991] and rules out an endogenous production shift towards specific crops that experience a rise in their global price. Finally, to ensure that out-migration and global food prices are not simultaneously determined by variation in third factors such as time-varying oil prices or global economic activity, we incorporate time fixed effects into our analyses.

Our findings complement the existing literature in a number of ways. First, by

analysing implications of distant effects of climatic events, we depart from the existing climate migration research that has thus far primarily focused on locally occurring climatic impacts as potential drivers of migration [Millock, 2015, Hoffmann et al., 2020, Sedova et al., 2021]. We show that an increase in locally-relevant global crop prices by one standard deviation increases household out-migration from agricultural households by about one third of the net effect of a comparable local weather shock. Global prices thus constitute an important driver of household out-migration, implying that the link between short-term variation in climate and labour migration from agricultural households has thus far been underestimated due to second-order effects. Second, we contribute to recent efforts to better understand the contextual effects of climate migration [Cattaneo et al., 2019]. We show that implications for migration of both global and local shocks caused by a changing climate similarly depend on the initial household wealth. Positive income shocks caused by higher international food prices help to relax the budget constraint of poor agricultural households and facilitate migration. The study thereby also contributes directly to the growing literature on the income elasticity of migration [Cattaneo and Peri, 2016, Hirvonen, 2016, Bazzi, 2017]. Instead of solely focussing on income effects brought about by shocks to the quantity produced by agricultural households, our study is the first to use variation in global crop prices as a plausibly exogenous transitory shifter of household income in Sub-Sahara Africa, the region with the largest share of households dependent on small-scale farming globally. With the exception of Bazzi [2017], who studies the income elasticity of international migration exploiting a persistent import ban on rice in Indonesia, prices of agricultural goods have thus far not been used as a source of exogenous variation in income to study the probability of migration along the household wealth distribution. We further add to this literature by showing that the source of the income shock agricultural households experience plays a role for destination choices among migrants, likely due to the general equilibrium effects on real income induced by a rise in agricultural prices. Finally, we further contribute to the literature on the link between climate, migration and conflict. It has been shown that, while climate-related migration can lead to conflict, climate-related conflict can also trigger migration [Ash and Obradovich, 2020, Abel et al., 2019, Missirian and Schlenker, 2017]. In this study, we add to this literature by considering the implications of conflict induced by global crop price fluctuations for household out-migration. We show that neither factor nor output conflict plays a role for this particular type of migration.

The remainder of this paper proceeds as follows. The next section presents the theoretical framework. In section 4.3, we provide an overview of the data and discuss our

constructed variables in detail. In section 4.4, we lay out our empirical strategy. Our findings are presented in section 4.5. In section 4.6, we extend the analysis to considering conflict as a potential driver of climate migration. The last section provides a discussion and concluding remarks.

4.2 Theoretical framework and research hypotheses

The framework we suggest to guide our thinking on the interdependence of global crop prices and the household migration decision closely follows Dustmann and Okatenko [2014] and incorporates insights from Bazzi [2017]. We use this parsimonious framework to build the intuition that (i) the effect of global food prices on household out-migration is a priori ambiguous on aggregate due to its differential effects along the household wealth distribution and (ii) is expected to differ between agricultural and non-agricultural households. We further compare income effects caused by locally-relevant global prices to those caused by the local quantity produced to show that these effects are qualitatively similar for agricultural households, but not for non-agricultural households.

4.2.1 General framework

In the simple framework we suggest in the following, households themselves are immobile and all decisions they make relate to sending one of their household members as a migrant. Similar to Dustmann and Okatenko [2014], we formalise this household decision as a comparison of utility flows in the current location compared to potential destinations. We use subscripts $l = h$ (home) and $l = d$ (destination) for all variables relating to location choices l of the potential migrant. The subscripts $y = a$ (agricultural household) and $y = n$ (non-agricultural households) describe the type of household y. The flow of utility in location l for household type y is then given by:

$$U_{hly} = INC_{hy}(p_{hy}, q_{hy}) + inc_{ly}(p_{ly}, q_{ly}) + \epsilon_{ly}, \tag{4.1}$$

where INC denotes the household real income generated by all non-migrant household members and inc denotes the real income generated by the potential migrant. ϵ_{ly} denotes a random variable capturing all non-income utility components. Both INC and inc are a function of the price of locally grown crops, p_l, and the local quantity produced, q_l. Both p_l and q_l are exogenously determined. For p_l, this is due to short-term fluctuations

in world market prices. For q_l, it is due to unpredictable fluctuations in local growing conditions. In addition, we assume

$$corr(q_h, p_h) = 0. \tag{4.2}$$

Thus, we rule out feedback loops between the local quantity produced at home and the exogenously determined global prices, an assumption we will discuss further in section 4. The second term of $inc_{ly}(p_{ly}, q_{ly})$ is a simplification: The income generated by the potential migrant is a future (expected) income flow; however, since no information on the future is available, households maximise their utility based on contemporarily observed income flows in all destinations d, which they assume to be accurate measures of what the potential migrant would earn in the future.

Migration is costly and when households make the decision to send a migrant, households are budget constrained by their initial household wealth. Assuming that households face borrowing constraints we can write this budget constraint as:

$$W_x + INC_{hy}(p_{hy}, q_{hy}) + inc_{hy}(p_{hy}, q_{hy}) \geq C_d, \tag{4.3}$$

where C_d is the location specific migration cost and W_x is the initial, idiosyncratic wealth of household x. $INC_{hy}(p_{hy}, q_{hy})$ is the income generated by the core household and $inc_{hy}(p_{hy}, q_{hy})$ is the income earned by the migrant at home.[1] We therefore implicitly assume that the household decision to send a migrant is made after income is earned at home. Equation 4.3 thus describes the threshold above which we could potentially observe migration from a given household. Combing equation 4.1 and 4.3 allows us to write the probability of a household sending a migrant as:

$$Pr(migration) = Pr(U_{hdy} > U_{hhy}, W_x + INC_{hy}(p_{hy}, q_{hy}) + inc_{hy}(p_{hy}, q_{hy}) \geq C_d). \tag{4.4}$$

Thus, two potential reasons may lead us to observe changes in household out-migration rates when households experience exogenous shocks to their household income: First, for non-budget constrained households, i.e. when the following equation holds

$$W_x + INC_{hy}(p_{hy}, q_{hy}) + inc_{hy}(p_{hy}, q_{hy}) \geq C_d,$$

[1] The assumption that borrowing constraints are a negative function of household wealth would lead to a qualitatively similar conclusion.

the opportunity costs of migration are altered in response to income shocks. A positive income shock increases U_h and renders staying home more attractive, while the reverse holds for a negative income shock. However, for households whose household budget lie marginally above C_d, a negative shock to $INC_{hy}(p_{hy}, q_{hy}) + inc_{hy}(p_{hy}, q_{hy})$ may push them below the budget constraint such that it becomes binding and out-migration rates decrease from these households.

Second, for households that are initially budget constrained such that

$$W_x + INC_{hy}(p_{hy}, q_{hy}) + inc_{hy}(p_{hy}, q_{hy}) < C_l,$$

a positive income shock may increase migration if the shock to $INC_{hy}(p_{hy}, q_{hy})+inc_{hy}(p_{hy}, q_{hy})$ is sufficiently large and $U_d > U_h$ still holds. This is, household out-migration only increases if the increase to U_h is not too large to make staying the relatively most attractive option. For budget-constrained households, negative income shocks have no effect on the decision to send a migrant since they simply remain below C_l.

4.2.2 Agricultural and non-agricultural households

In this subsection, we turn more closely to the household income, given by

$$INC_{hy}(p_{hy}, q_{hy}) + inc_{hy}(p_{hy}, q_{hy}),$$

and its dependence on globally determined prices and the growing conditions of locally produced crops by type of household. For the remainder of this section, we assume that both $INC_{h,y}$ and $inc_{h,y}$ are differentiable in p_h and q_h. The sign of the first derivative then depends on the type of household, y. We expect changes in global food prices of locally grown crops p_h to have a strictly non-negative effect on the household income of agricultural households, a, which we define as net producers. All other crop prices equal, only subsistence farming households that consume all their produce do not experience a positive income shock to their household wealth. Thus, we assume the following derivatives:

$$\frac{\partial inc_{ha}}{\partial p_{ha}} \geq 0; \frac{\partial INC_{ha}}{\partial p_{ha}} \geq 0. \tag{4.5}$$

Similarly, a positive (negative) shock to local growing conditions that increases (de-

creases) the quantity produced, increases the income of agricultural households:

$$\frac{\partial inc_{ha}}{\partial q_{ha}} \geq 0; \frac{\partial INC_{ha}}{\partial q_{ha}} \geq 0. \tag{4.6}$$

Note that in some cases, households farming at subsistence levels may consume all additional produce such that the derivative in equation is not strictly larger than zero.

For non-agricultural households, which we define as consumers of agricultural goods, this relation is unambiguously non-positive. If locally consumed and produced crop varieties partly coincide, i.e. if local crop consumption patterns are partly correlated with locally produced crops, real income - and thus household wealth - declines for non-agricultural households. If the correlation between local consumption patterns and local production equals zero, we would observe no effect of global prices relevant for local production on the household budget of non-agricultural households. Thus, for non-agricultural households, we have

$$\frac{\partial inc_{hn}}{\partial p_{ha}} \leq 0; \frac{\partial INC_{hn}}{\partial p_{ha}} \leq 0. \tag{4.7}$$

We further expect non-agricultural households to be significantly less affected by changes in local growing conditions. With their household income not directly related to the local quantity produced and local consumer prices following the world market, we expect real income of non-agricultural households to remain unaffected by the quantity of crop harvested locally. Thus, we assume that

$$\frac{\partial inc_{hn}}{\partial q_{ha}} = 0; \frac{\partial INC_{hn}}{\partial q_{ha}} = 0. \tag{4.8}$$

In summary, we derive the following research hypotheses from the above simple framework: The aggregate observed effect of exogenously determined fluctuations in household income on household out-migration is a priori ambiguous. It depends on two factors. First, it depends on the initial wealth distribution of households through the interplay of three forces: the opportunity costs of migration, the household budget constraint and the migration costs. Second, the aggregate effect depends on the type of household, which determines the direction of the wealth shock induced by global prices and the quantity produced locally: Agricultural households can be expected to experience an income (and thus, wealth) increase when locally-relevant global crop prices rise or local growing conditions improve, whereas non-agricultural households experience an income shock that is

unambiguously non-positive for global prices or strictly zero for local growing conditions.

In the following, we describe in detail how we test the derived hypotheses empirically.

4.3 Data

To create our dataset, we draw on several data sources. The household data are presented in section 4.3.1. Data used to generate the producer prices are discussed in section 4.3.2. Our variables related to weather data are presented in section 4.3.3. In section 4.3.4, we introduce the conflict data.

4.3.1 Household data

To generate our main dependent variable and the various derivatives of it, we draw on the World Bank's African Migration and Remittances Surveys (AMRS).

Within AMRS, households were surveyed in five countries in Sub-Saharan Africa. In Kenya, Burkina Faso, Nigeria and Senegal interviews were conducted in late 2009, in Uganda in early 2010.[2] AMRS contains retrospective information on non-resident household members' years of out-migration as well as their destination choices and their reasons for moving. In sum, this information is available for approximately 2000 households in each country. We draw on these household-specific migration histories to generate the dependent variable. To minimize the errors related to the retrospection, we limit our sample to the ten years prior to the year of the interview to generate a nine-year household time series from 2000 to 2008.[3] We restrict our sample to households whose head is 25 years or older in 2000 to account for the fact that households with heads younger than 25 years old were unlikely to exist in 2000 [Gray and Wise, 2016]. We further include return migrants, defined as migrants who left the household in the past and returned at a later stage; however, these constitute less than 5 per cent of our migrant sample. Household members that left for the purpose of studying are not treated as migrants to account for the fact that migration for education reasons is guided by different dynamics than labour migration, which is the primary focus of our analysis. Our main dependent variable (M) is binary and takes on a value of one if the sum of households' (h) migrants (m) in a given year (t) increases compared to the preceding year and is equal to zero otherwise. More

[2] We treat all countries in our sample as being interviewed in 2009 for consistency reasons.

[3] Using the AMRS data, Gray and Wise [2016] apply a similar approach to generate their migration variables for a six-year migration panel.

formally, we define

$$M_{ht} = \begin{cases} 1, & \text{if } m_{ht} > m_{ht-1} \\ 0, & \text{otherwise.} \end{cases}$$

To test our hypotheses as presented in section 4.2, we split our sample into households whose livelihoods do (throughout the text referred to as agricultural) and do not (throughout the text referred to as non-agricultural) depend on agricultural production. Households are considered to depend on agriculture if they own agricultural land or at least one of their members is full-time employed in agriculture. Non-agricultural households, on the contrary do not own land and none of their members works in agriculture. Districts are the finest geographical level on which we are able to reliably identify our households.

Table 4.1 presents the corresponding summary statistics where N indicates the number of household-years. The table shows that almost 8% of agricultural and almost 7% of non-agricultural households experience out-migration of at least one household member in every given year. Most of household out-migration takes place internally to urban areas for both types of households. International migrants from agricultural households are likely to move to other African countries, while international migrants from non-agricultural households are likely to move to OECD countries.

Variable	Mean	Std. Dev.	Min.	Max.	N
Agricultural households					
Overall	0.076	0.265	0	1	52101
Internal	0.058	0.235	0	1	52101
Internal: Rural	0.013	0.113	0	1	52101
Internal: Urban	0.046	0.209	0	1	52101
Africa	0.02	0.141	0	1	52101
OECD	0.008	0.091	0	1	52101
Non-agricultural households					
Overall	0.066	0.248	0	1	17307
Internal	0.046	0.209	0	1	17307
Internal: Rural	0.007	0.082	0	1	17307
Internal: Urban	0.04	0.195	0	1	17307
Africa	0.009	0.093	0	1	17307
OECD	0.019	0.135	0	1	17307

All variables were constructed at the household-year level using World Bank's African Migration and Remittances Surveys data. They are binary and take on a value of one if a household increased its number of out-migrants to a given destination compared to the year before, and zero otherwise.

Table 4.1: Summary statistics: Households' migration (by destination)

To proxy household wealth in the cross-section, we construct a household wealth index

similar to Dustmann and Okatenko [2014]. The index is based on the following seven survey questions:

- *Does the household own the house it lives in?* Yes/No.
- *Does the household have access to electricity?* Yes/No.
- *Does the household have access to piped water?* Yes/No.
- *Does the household own a television?* Yes/No.
- *Does the household own a computer?* Yes/No.
- *Does the household own a bank account that was not set up in response to a migrant leaving the household?* Yes/No.
- *Has the head of household attended a school?* Yes/No.

We conduct a principal component analysis on these variables and use the factor loadings of the first principle component as weights to construct an aggregate wealth index. The corresponding Kaiser-Meyer-Olkin measure of sampling adequacy indicates a value of 0.74, supporting the suitability of the approach [Dziuban and Shirkey, 1974]. The resulting wealth index is then normalised to lie between 0 and 1.

Since the variables which we use to construct the wealth index can only be observed in 2009, they can be endogenous to the previous migration decision. We therefore also approximate wealth by a wealth measure based only on variables with information for the year 2000, the first year in our household panel. The first variable we use to calculate the pre-migration wealth is the indicator on whether or not the household head received any kind of formal education, guided by the idea that the level of education of the household head is a good predictor of household wealth in African countries [Duflo, 2001, 2004, Maccini and Yang, 2009, Wantchekon et al., 2015]. We implicitly assume that the decision to receive formal education is finalised at the beginning of the observation period and does not change over time.[4] The second variable is based on the question whether the respective household - this is, anyone in the household - owns a bank account. The question is followed up by a second question on whether this bank account was opened in response to a previous member of household leaving the household, allowing us to correct the bank account indicator for reverse causality. We then combine these two survey questions to construct a parsimonious wealth measure exogenous to the household decision to send a

[4]Since we limit our sample to households whose head is 25 years or older at the beginning of the observation period, the assumption of schooling being finalised is plausible.

migrant. This measure divides households into the following three categories: low wealth households (0) with no bank account and where the head has zero years of schooling, medium wealth households (1) with either a bank account or where the head has received some schooling and upper-wealth households (2) with both a bank account and where the head attended school. Table 4.2 presents the corresponding summary statistics.

Variable	Mean	Std. Dev.	Min.	Max.	N
Owns house	0.716	0.451	0	1	7712
Access to electricity	0.482	0.5	0	1	7712
Access to piped water or public well	0.718	0.45	0	1	7712
Attended school (head)	0.569	0.495	0	1	7712
Owns computer	0.116	0.32	0	1	7712
Owns television	0.471	0.499	0	1	7712
Bank account (pre-migration)	0.335	0.472	0	1	7712
Wealth index	0.406	0.297	0	1	7712
Wealth index pre-migration	0.903	0.813	0	2	7712

All wealth index variables were constructed at the household level using World Bank's African Migration and Remittances Surveys data. The variable *Wealth index* is a binary indicator constructed using household level information from 2009 as shown in the upper part of the table, with higher valuer indicating more wealth. *Wealth index pre-migration* uses information from 2000. It is a categorical variable dividing households into low (0), medium (1) and upper-wealth (2) categories.

Table 4.2: Summary statistics: Households' wealth indexes

4.3.2 Global prices

We follow McGuirk and Burke [2020] in the construction of a plausibly exogenous price index that allows us to analyse the causal effect of price changes on migration. Similar to the authors, we require price data that varies sufficiently over time, is not determined by local factors and that allows us to differentiate real income effects across households. We therefore generate district-specific price time series by combining exogenous *temporal* variation in global crop prices with local *spatial* variation in crop production at the beginning of our observation period.

(Producer) Price Index (PPI): To generate the spatial variation in the PPI, we utilise the high-resolution crop-specific fraction of harvested area in year 2000 (i.e. the first year of our observation period) which contains information on harvested area and yield for 175 crops, initially compiled by Monfreda et al. [2008]. The authors create this land use dataset by combining national-, state- and county-level census statistics with a global dataset of croplands with a 5×5 minute grid cell resolution. Using these data, in Figures 4.8, 4.9, 4.10 and 4.11 in Appendix 4.B, we provide illustrative examples of how production of different commodities differs by country. To generate the temporal variation in the

PPI, we draw on annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. Prices are indexed at 100 in year 2010. We then compute the annual district-specific PPI by combining the temporal variation of commodity prices and the spatial variation of crop-specific fraction of harvested area in 2000 in the following way:

$$PPI_{dt} = \sum_{i=1}^{n} (P_{it} \times F_{idc}), \tag{4.9}$$

whereby crops $(i...n)$ capture a set of 12 major traded crops for the five countries in our dataset that are simultaneously covered by the land use dataset and for which international prices exist, F_{idc} captures the district-specific crop share of land. For a full list of crops used to generate the international food prices, see Table 4.10 in Appendix 4.B. To better capture the nature of unexpected shocks, we express the PPI as a percentage change from its district-specific long-run mean constructed for the pre-sample period 1990-1999. To summarise, the district-level variation of PPI comes from annual global crop price changes and a district-specific mix of locally produced crops. In addition, following McGuirk and Burke [2020], we also generate two disaggregated versions of the PPI for robustness checks: i) PPI (food), which captures price index for crops that constitute more than 1% of calorie consumption in the overall sample as suggested by food consumption data from the UN Food and Agriculture Organization (FAO), and ii) PPI (cash) which is the price index covering the remaining crops.

Figure 4.6 in Appendix 4.B further captures how the PPI developed over time, suggesting a sharp spike during years of the food price crisis in 2007/08 in all countries. The spatial distribution of average PPI for the years of 2007/08 food price crisis in Figure 4.7 in Appendix 4.B suggests that, likely due to the spatial correlation of soil-suitability in combination with spatially correlated climatic conditions that result in geographically correlated crop-production, our price index shows patterns of spatial correlation. We will attend to this phenomenon in more detail in section 4.5.3.

One of the key identifying assumption in all subsequent analyses - and in fact, of all empirical studies utilising local weather as a source of exogenous variation in local agricultural income - is the exogeneity of global crop prices to local production in Sub-Saharan Africa. If the quantity of agricultural goods produced locally was a predictor of global commodity prices, these prices would, to some extent, co-move with local production and attenuate shocks on quantities produced locally. A number of reasons should convince

the reader that agricultural households in Sub-Saharan Africa are indeed price takers: First, to isolate the effect of global crop prices from local weather conditions, we focus on the observation period from 2000 to 2008. This time span covers the global food price crisis 2007/08, when global food prices rose sharply for reasons entirely exogenous to agricultural activity in Sub-Sahara African [Berazneva and Lee, 2013, Demeke et al., 2009, Dorward, 2012]. Second, between 1989-2013, the entire continent of Africa accounted for only 5.9% of global cereal production, minimising its effect on global prices [McGuirk and Burke, 2020]. Finally, we show in all our analyses that the effect of global prices on our outcomes are not sensitive to conditioning estimates on the local quantity produced, proxied by local weather conditions.

4.3.3 Local weather

To generate a set of climate-related variables, we draw on ERA5 reanalysis data produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) [Copernicus Climate Change Service (C3S), 2017]. ERA5 is the fifth generation of ECMWF atmospheric reanalyses of the global climate. It is a high quality reanalysis dataset which relies on information from weather stations, satellites, and sondes. ERA5 provides data at a geographical resolution of 31km and has been regridded to a 0.25×0.25 degrees latitude-longitude grid. Currently, the weather data is available from January 1979 with a temporal resolution of up to one hour. We use the daily mean temperature as well as total monthly precipitation and aggregate these to the district level, using Google Earth Engine.

The existing literature shows that the effect of temperature on economic outcomes is highly non-linear [Burke et al., 2015, Carleton and Hsiang, 2016, Schlenker and Roberts, 2009, Kalkuhl and Wenz, 2020]. Growing degree days (GDD) is one common way to capture this non-linear relationship. One degree day counts the total amount of degrees above a lower threshold as long as the mean local temperature is below an upper threshold on a given day. If mean temperature (T) exceeds the upper threshold, degree days are counted as the difference between the upper bound and a lower bound. More formally, growing degree days D above a lower threshold l_1 and below an upper threshold l_2 are

defined as:

$$D = \begin{cases} l_2 - l_1 & \text{if } T > l_2 \\ t - l_1 & \text{if } l_1 < T \leq l_2 \\ 0 & \text{if } T \leq l_1. \end{cases}$$

GDD then capture the total number of degree days over the crop-growing season, defined as the time period stretching from June to August (JJA) [Dell et al., 2014, Schlenker et al., 2007, Schlenker and Roberts, 2009].

Following the literature, we generate two GDD-related variables at the district-level: i) degree days between 10 and 30°C, and ii) degree days above 30°C. The intuition behind the choice of the bounds is that temperature between 10-30°C generally enhances yield, while temperatures above 30°C is considered yield-decreasing [Schlenker and Roberts, 2009, Schauberger et al., 2017]. We control for the average growing season precipitation, measured as the average of daily (total) precipitation during the growing season in meters height collected on each square meter. However, precipitation is not of the main interest in our analysis for several reasons. First, in the context of Sub-Saharan Africa, precipitation has been shown not to be an important predictor of migration [Missirian and Schlenker, 2017] and conflict [Burke et al., 2009]. Second, relative temperature changes under future climate change scenarios translate into much larger changes in yields than do precipitation changes in Sub-Saharan Africa [Schlenker and Lobell, 2010]. Third, even though weather data sets agree on long-run averages, particularly in the case of precipitation they do not necessarily agree on anomalies [Auffhammer et al., 2013]. Since deviations from the mean are the main source of identification in our setting, this could potentially be problematic. We nevertheless control for precipitation in all regression analyses for completeness. Table 4.3 presents the summary statistics of main local climatic variables of interest at the district-level.

Variable	Mean	Std. Dev.	Min.	Max.	Units
Growing Degree Days (10-30°C) (GDD1030)	13.993	2.844	4.889	17.722	Hundred
Growing Degree Days (>30°C) (GDD30)	0.299	0.492	0	3.1	Hundred
Precip. (JJA)	0.121	0.073	0.008	0.45	m
N		1260			

All weather variables were constructed at the district-year level using ERA5 reanalysis data and capture conditions during the growing season covering months June-August.

Table 4.3: Summary statistics: Climate-related variables at the district-year level

4.3.4 Conflict data

We follow state-of-the-art approaches in conflict analyses [McGuirk and Burke, 2020, De Winne and Peersman, 2019] and distinguish between i) small-scale *output* conflicts associated with appropriation of surplus, and ii) large-scale *factor* conflicts over the control of territory. We only focus on conflicts in the second half of each year (i.e. the period from July to December) to link all conflict events to the district-specific crop yield of each growing season from June to August. Mid-growing season, households will have plausibly assessed their potential agricultural income.

Output conflicts: To capture output conflict, we draw on the Armed Conflict Location and Event Data Project (ACLED) database [Raleigh et al., 2010]. ACLED provides temporally and geographically disaggregated data on dates, actors, locations, fatalities, and types of all reported political violence and protest events, collected from a range of media and agency sources. Since output conflict is likely to be transitory and disorganised, we further draw on information on riots, protests and violence against civilians [McGuirk and Burke, 2020, De Winne and Peersman, 2019]. We then construct a binary output conflict variable on the external margin of output conflict incident, covering our sampling period from 2000 to 2008.[5]

Factor conflicts: We further draw on geo-coded conflict-related fatality count data from the Uppsala Conflict Data Program (UCDP) [Sundberg and Melander, 2013, Pettersson and Öberg, 2020]. UCDP provides temporally and geographically disaggregated information on conflict events, which entail the use of armed force by an organised actor against another organised actor, or against civilians, resulting in at least one direct death. It covers all dyads that have crossed a threshold of 25 battle deaths per year. The data is gathered from local and national media, agencies, NGOs and international organisations. Since UCDP data capture relatively larger scale conflicts, the data is suitable to approximate conflicts associated with the control of territory, i.e. factor conflicts [McGuirk and Burke, 2020]. We aggregate the fatality counts to the district-year level for the period 2000-2008 to match the fatality counts with our other data sources. Similar to output conflict, our main variable of interest is binary and takes the value one if any conflict event took place in a given district-year between July to December, and zero otherwise.

The constructed conflict-related variables at the district level are summarised in Table

[5]The decision to code our conflict variables as binary (and thus only consider the external margin) is guided by the conflict literature [Bazzi and Blattman, 2014, Berman et al., 2017, Nunn and Qian, 2014, McGuirk and Burke, 2020] and reduces the potential measurement error stemming from the recording of the original conflict events [McGuirk and Burke, 2020]

4.4 below.

Variable	Mean	Std. Dev.	Min.	Max.
Output conflict (July - December)	0.2405	0.4275	0	1
Factor conflict (July - December)	0.0722	0.259	0	1
N		1260		

Output conflict was constructed using Armed Conflict Location and Event Data Project data and captures occurrence of smaller-scale conflicts at the district-year level. Factor conflict was constructed using Uppsala Conflict Data Program data and captures occurrence of large-scale conflicts at the district-year level.

Table 4.4: Summary statistics: Conflict occurrence at the district-year level

4.4 Methodology

We first examine the effect of the exogenous variation in international food prices measured by the PPI on households' decision to send out a migrant, while controlling for local climatic variables. To do so, we estimate the following baseline equation:

$$Pr[M_{hdt} = 1|x_{dt}, \phi_t, \alpha_h] = \beta_0 + \beta_1 PPI_{dt} + \beta_2 GDD_{dt;10:30} + \beta_3 GDD_{dt;30:\infty} + \beta_4 GP_{dt} + \alpha_h + \phi_t + \epsilon_{hdt}. \quad (4.10)$$

Thus, our binary indicator capturing a household-specific (h) increase in migration in a given year (t) relative to the preceding year (M_{hdt}) is regressed on yearly district-specific (d) food price index (PPI), district-specific number of growing degree days (GDD) between 10-30°C and above 30°C, district-specific average precipitation during the growing season and its squared term (summarised by GP) and year (ϕ_t) and household (α_h) fixed effects. By applying a two-way fixed effects approach, the identification comes via deviation of global prices from their historical district-specific mean over time. Thus, a given households' out-migration rate is compared under different price regimes. The year fixed effects ensure that out-migration and global food prices are not simultaneously determined by variation in third factors, such that the estimated effect can be interpreted causally. Figure 4.1 presents the yearly, district-specific variation in PPI from the district-specific mean from 2000 to 2008, which serves as the main source of identification of the response coefficient. We cluster standard errors at the level of the treatment, i.e. at the district-level.

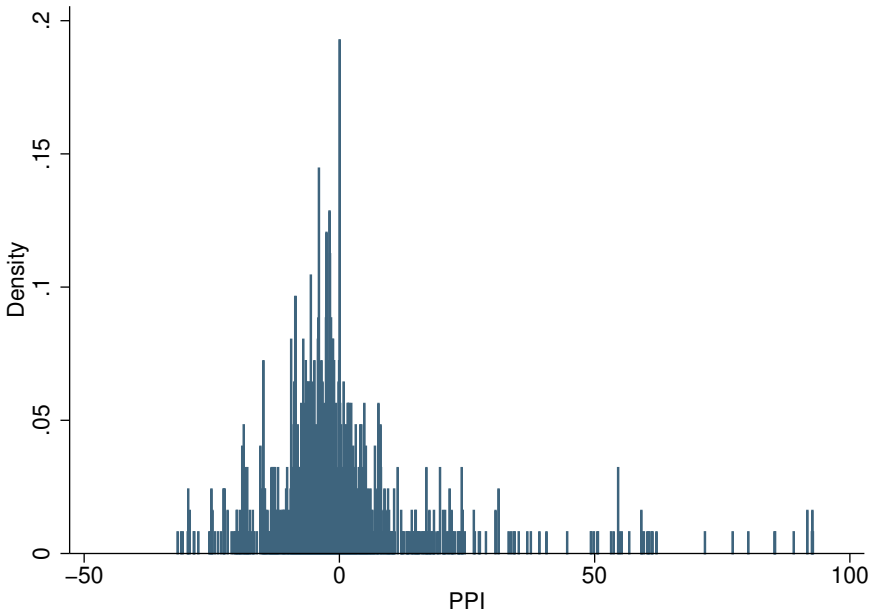


Figure 4.1: Within-district variation in PPI

In our next specification, we then study how international prices interact with households' wealth and affect households' migration decision in the following way:

$$Pr[M_{hdt} = 1|x_{dt}, \phi_t, \alpha_h] = \beta_0 + \beta_1 PPI_{dt} + \sigma PPI_{dt} \times Wealth_h + \beta_2 GDD_{dt;10:30} + \beta_3 GDD_{dt;30:\infty} + \beta_4 GP_{dt} + \alpha_h + \phi_t + \epsilon_{hdt}. \quad (4.11)$$

Thus, the coefficients on β_1 and σ combined capture the differential effect of locally-relevant global prices along the household wealth distribution. $Wealth_h$ enters the equation as either a continuous (post-migration wealth) or categorical (pre-migration wealth) variable as explained in the previous section.

We estimate equations 4.10 and 4.11 using both a reduced-form linear probability model (LPM) and a logistic (Logit) model. The LPM assumes a linear relation between the household decision to send a migrant and the changing local income conditions. While no additional modelling choices are required, this assumption is potentially strong in a setting when the majority of household-years contain zero values. Maximum likelihood based probability models such as the Logit allow for a more flexible, non-linear probability function more suitable for such a setting and recent advances in logistic models that incorporate large number of fixed effects also overcome the incidental parameter problem inherent to these models (see [Lancaster, 2000] for a detailed discussion).

However, maximum-likelihood based fixed effects Logit models also have their disad-

vantages in the setting at hand: First, since it is mathematically impossible for maximum-likelihood models to converge when there is no within-category variation, these models often fail to converge when high dimensional fixed effects are incorporated. For example, when including country-by-year fixed effects, the likelihood function fails to find a global maximum for all subsamples. Second, when reporting the marginal effects from the Logit models, additional assumptions on the values of the fixed effects -which are not estimated - are required.[6]

We therefore suggest an estimation strategy based on an LPM with a common time-trend and climatic controls to obtain our parameters of interest as our preferred specification. To ensure that the choice of the model does not drive our results, we also show obtained average marginal effects from the conditional fixed effects Logit specification following Kitazawa [2012] and Kemp et al. [2016]. We then also estimate a more conservative specification of our models that include country by year fixed effects (LPM) or country-specific linear time trends (Logit). We do not choose these specification as our preferred ones for two main reasons. First, because of within-country spatial correlations in our main variable of interest - which we will explore further in subsection 4.5.3 - the country-year trends could potentially absorb a significant share of the variation of interest. Second, the additional loss in degrees of freedom is critical compared to the additional precision our estimates gain in a setting where the essential household fixed effects absorb a high number of degrees of freedom. Nevertheless, we will present these estimates for robustness.

As outlined in the introduction, we further aim to examine conflict as a potential mechanism that could in part explain the relationship between producer prices and household out-migration. The analysis of the channeling effect of conflict in similar econometric settings is a widely discussed empirical challenge (see for example Berlemann and Steinhardt [2017]). Including conflict as a control variable in 4.10 could bias the coefficient on our main independent variables if conflict itself is an outcome of changes in locally-relevant food prices and local climatic conditions, a problem commonly referred to as an *over controlling* [Dell et al., 2014] or a *bad control* problem [Angrist and Pischke, 2008].

We therefore structure our thinking on the global prices-conflict-migration nexus as follows. We start out by estimating the association of household out-migration and the different types of conflict in the following parsimonious regression framework:

[6]Some modellers in the migration literature choose to report the marginal effects by setting the fixed effects to zero (see for example [Bazzi, 2017]. However, this is an arbitrary choice.

$$Pr[M_{hdt} = 1|x_{dt}, \phi_t, \alpha_h] = \theta_0 + \theta_1 Conflict_{dt} + \alpha_h + \phi_t + \epsilon_{hdt} \tag{4.12}$$

where the dependent variable is defined as the household out-migration from household h, residing in district d in year t as before. We distinguish between output and factor conflict, both of which enter equation 4.12 separately as $Conflict_{dt}$. The estimated coefficient on conflict, θ_1, should not be interpreted causally due to the potential problems of reverse causality and omitted variable bias. Nevertheless, these correlations can provide us with first valuable information as they reveal how household out-migration and conflict relate to one another. To more directly explore the link between locally-relevant global prices and conflict, we study the district-level association between the PPI and the different types of conflict in Appendix 4.F.

4.5 Results

In this section, we present all outcomes from our regression analyses. Specifically, in section 4.5.1, we present results of the aggregate association between global food prices and migration. In section 4.5.2, we study the heterogeneity of these effects along the initial household wealth distribution. Finally, in section 4.5.3, we turn to destination choices in response to changes in locally-relevant global food prices. We present outcomes from both the LPM and the Logit model in all of our main analyses and for the more specific results, we only show our preferred specification as described in the previous section.[7]

4.5.1 Aggregate effect of global food prices

Tables 4.5 and 4.6 show the estimated effects of changes in global food prices (PPI) on the probability of migration for agricultural and non-agricultural households respectively. We will refer to model (5) as the result from our preferred specification.

For the **sub-sample of agricultural households**, both the LPM and the Logit show similar outcomes throughout different specifications. Our preferred specification (5) suggests that a one percentage point increase in locally-relevant global food prices over its district specific long-run mean increases the likelihood that households send out a migrant

[7]Note that output tables of the Logit regressions show the average semi-elasticity of $Pr[M_{hdt} = 1|x_{dt}, \phi_t, \alpha_h]$ with respect to our variables of interest. The magnitude of these coefficients is therefore not directly comparable to the OLS estimates, which capture level-level (percentage point) changes.

by 0.06 percentage points. Even in the most demanding model specifications (3 and 6), the effect remains highly visible and does not change in its order of magnitude. Importantly, comparing models (1) to (2) and (4) to (5), it becomes evident that the magnitude of the effect of producer prices on household out-migration remains unchanged when conditioned on local weather. These controls even add precision to the obtained estimates, albeit only marginally. This first finding provides reassurance that international food prices and local weather are not linked causally.[8]

The estimated coefficients on the local weather variables indicate that an increase of the number of degree days during the growing season above 30°C by 100 decreases the likelihood of households sending one of their members as a migrant in the same year by almost three percentage points (model 5). On the other hand, 100 additional growing degree days between 10 and 30°C significantly increase the likelihood of households sending out a migrant by 2 percentage points. A note of caution is warranted when interpreting these findings. While the sign of the estimates on our generated weather variables is stable throughout all specifications, the coefficients are not always significant at conventional statistical levels. A potential reason for this is the inherent collinearity of the two variables at their intersection points around the bounds defined in the previous section; however, the particular modelling of the variables is necessary to capture the non-linearity in increasing temperatures as discussed in 4.3.3.

For the **sub-sample of non-agricultural households**, we generally do not find any statistically significant effects of global food prices on household out-migration. However, the sign of the association is positive throughout specifications and turns significant at the five percent level in model (3); these outcomes could capture the pull effects of the agricultural sector, a potential explanation we will offer more detail on in section 4.5.3 and Appendix 4.E. We do not find a significant impact of adverse local weather on non-agricultural households. This suggests that if local production conditions deteriorate, non-agricultural households are able to diversify their consumption via world markets, as implied by equation 4.8.

Taken together, for households whose livelihoods depend on agricultural production,

[8]To further test that PPI is exogenous and thus is not determined by local conditions we regressed it on local climatic variables. By using a fixed effects panel data regression, location and time effects absorb potential large-scale correlation of climatic events and related trends. The remaining variation identifies responses from deviations of local climatic conditions over time from the long-term, location-specific mean. The results suggest that local climate-related variations do not significantly predict variation in PPI. Since we use the same identification strategy in the main analysis, we can plausibly claim that the variation that identifies PPI responses is not defined by local conditions as potential sources of correlation are captured by fixed effects. The estimation results are available upon request.

we find that higher international food prices facilitate migration, an indication that positive producer price shocks have a similar migration-inducing effect as positive weather shocks. Our explanation as proposed by the theory in section 4.2 is that positive income shocks, either via increases in the PPI or yield enhancing temperatures, push households above the previously binding budget constraint. Lower PPI or yield decreasing temperatures, on the contrary, reduce migration by imposing a stricter budget constraint [Cattaneo and Peri, 2016, Deaton, 1989, Bellemare et al., 2018, McGuirk and Burke, 2020]. Households in the agricultural context of Sub-Saharan Africa are typically characterised by low income levels [McGuirk and Burke, 2020]. Thus, the migration-inducing effect of the PPI appears to be larger than the rising opportunity costs of out-migration due to better income opportunities at home. We turn to these suggested interpretations of our aggregate findings in more detail in the next section.

To better understand the estimated associations, we further run the fully specified LPM model by distinguishing the PPI of cash and food crops (see Table 4.11, Appendix 4.C). Finally, we also run the fully specified LPM models with common and state-specific trends, presented in Appendix 4.C, to check our results for robustness with respect to alternating definitions of the growing season. In Tables 4.12 and 4.13 we present the outcomes for agricultural and non-agricultural households respectively. Even though none of these robustness checks seem to change our main results, we find that these results are mainly driven by price changes of food crops rather than cash crops, further suggesting that domestic consumption plays an important role in the setting of Sub-Saharan Africa. Moreover, the sensitivity tests confirm that households respond most strongly and consistently to the growing season conditions as defined in the main analysis.

	(1)	(2)	(3)	(4)	(5)	(6)
PPI	0.0051*	0.0052**	0.0106***	0.0006***	0.0006***	0.0004*
	(0.0026)	(0.0025)	(0.0025)	(0.0002)	(0.0002)	(0.0002)
DD1030		0.1800	0.0885		0.0219*	0.0203
		(0.1703)	(0.1493)		(0.0112)	(0.0135)
DD30		-0.3563	-0.2043		-0.0254*	-0.0448***
		(0.2322)	(0.2007)		(0.0148)	(0.0167)
N	23742	23742	23742	52101	52101	52101
R^2				0.012	0.013	0.017
Time trend	Year	Year	Country x Year	Year	Year	Country x Year
Model	Logit	Logit	Logit	LPM	LPM	LPM
Precip. controls	No	Yes	Yes	No	Yes	Yes

The dependent variable is binary and captures household-level out-migration incidence in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. The sample captures agricultural households only. Models 2-3 and 5-6 further control for growing season precipitation and their squared terms. Models 1-3 are estimated with fixed effects logit model and models 4-6 with LPM. Models 1-2 and 4-5 use a common and models 3 and 6 country-specific time trend. Model 5 corresponds to the preferred specification. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.5: Aggregate effect of PPI on the probability of migration: Agricultural households

	(1)	(2)	(3)	(4)	(5)	(6)
PPI	0.0039	0.0037	0.0077**	0.0004	0.0004	0.0003
	(0.0031)	(0.0031)	(0.0034)	(0.0003)	(0.0003)	(0.0004)
DD1030		0.2265	0.1836		0.0156*	0.0208*
		(0.1378)	(0.1380)		(0.0093)	(0.0113)
DD30		0.3363	0.5591		0.0205	0.0031
		(0.3793)	(0.3751)		(0.0214)	(0.0201)
N	7443	7443	7443	17307	17307	17307
R^2				0.016	0.016	0.023
Time trend	Year	Year	Country x Year	Year	Year	Country x Year
Model	Logit	Logit	Logit	LPM	LPM	LPM
Precip. controls	No	Yes	Yes	No	Yes	Yes

The dependent variable is binary and captures household-level out-migration incidence in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. The sample captures non-agricultural households only. Models 2-3 and 5-6 further control for growing season precipitation and their squared terms. Models 1-3 are estimated with fixed effects logit model and models 4-6 with LPM. Models 1-2 and 4-5 use a common and models 3 and 6 country-specific time trend. Model 5 corresponds to the preferred specification. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.6: Aggregate effect of PPI on the probability of migration: Non-agricultural households

Finally, to put the effect of locally-relevant prices into perspective, it is useful to compare changes in the PPI to the effect of local weather directly. To do so, consider

a hypothetical scenario where both the PPI and the DD30 increase by one standard deviation over their long-run mean. A back-of-the-envelope calculation that considers the non-linear effect of local temperatures on household out-migration (i.e. that accurately factors in the corresponding changes in DD1030), then shows that the overall climatic effect on migration of both, global prices and local temperature is positive. More precisely, the standardised effect of a global price increase on household out-migration is around 37% of the standardised (and so comparable) net effect of a rise in local temperature. Overall, these findings therefore suggest that in the context of Sub-Saharan Africa, the magnitude of short-term climatic effects on migration have thus-far been underestimated.[9]

4.5.2 The role of household wealth

In this section, we explore the role of households wealth, one of the contextual factors that could determine the direct associations between PPI and migration. To explore this particular heterogeneity, we turn to the results of regression model 4.11, detailed in section 4.4. First, in Table 4.7 we interact the PPI with the continuous wealth index and present the results from a series of fully specified LPMs and Logit models with and without country-specific trends for agricultural (models 1-4) and non-agricultural (models 5-8) households. Second, because the continuous wealth index is based on values partially measured in 2009 and the wealth it captures could therefore be endogenous to the migration decision, in Appendix 4.D in Tables 4.14 and 4.15, we interact the PPI with the exogenous categorical wealth index based on values from 2000 for agricultural and non-agricultural households respectively. Using this more exogenous measure without time variation also solves the over-controlling problem [Dell et al., 2014].

For the **sub-sample of agricultural households**, in Table 4.7 we find robust evidence across all specifications that wealthier households are less likely to send out migrants when the locally-relevant global prices increase. The marginal effects of PPI by wealth, i.e. the outcomes from the main specification (model 3), are depicted in Figure 4.2. The marginal effect is positive but decreases with increasing wealth. It remains statistically significant only for approximately the lower half of the wealth distribution. In Table 4.14 (in Appendix 4.D), we draw on the exogenous measure of wealth to test the validity of these results. In five out of six specifications, we find further evidence that richer households are less likely to send out migrants when the PPI increases. In our preferred model 5, the interactions show that if PPI increases by one p.p., medium-wealth and upper-wealth

[9]These calculations are available upon request.

households become 0.03 p.p. and 0.09 p.p. less likely to send out migrants respectively, compared to households with low levels of wealth. Figure 4.3 further presents these outcomes visually only for the lower and upper wealth categories (i.e. the medium wealth households are not included). It shows that a higher PPI increases the marginal probability of low-wealth and decreases the probability of upper-wealth households to send out migrants. When using this more exogenous wealth measure, the effect of the PPI is statistically significant from zero along the entire wealth distribution. The presented evidence underlines the interpretation of the direct effects of locally-relevant global prices on household out-migration, as suggested in the previous section. Implications of global price changes for migration differ depending on households' wealth. The findings strongly suggest that increases in household income induced by exogenous changes in relevant agricultural commodity prices push poor households above the previously binding budget constraint and facilitate their out-migration, exceeding contrary opportunity costs effect. The opportunity costs only start to play a more significant role for wealthier households, reducing their likelihood to send out migrant as income increases.

We do not find similar evidence for **sub-sample of non-agricultural households**. Both the baseline effect of the PPI and its interaction with household wealth show no statistically significant association with household out-migration, visible in both Table 4.7 and 4.15 (in Appendix 4.D). We interpret this as evidence that PPI does not capture prices of locally consumed crops, but rather captures the conditions in the agricultural sector. Thus, the weak evidence of positive aggregate effects of the PPI on out-migration from non-agricultural, net-consuming households shown in 4.5.1 seem to rather capture pull effects. We will discuss these in more detail in the next section.

Table 4.7: Heterogeneous effects of PPI by household wealth: Agricultural (models 1-4) and non-agricultural (models 5-8) households

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PPI	0.0069**	0.0128***	0.0008***	0.0006**	0.0033	0.0087	0.0005	0.0002
	(0.0027)	(0.0024)	(0.0002)	(0.0002)	(0.0081)	(0.0083)	(0.0007)	(0.0007)
PPI × Wealth	-0.0078*	-0.0106***	-0.0007*	-0.0005*	0.0007	-0.0018	-0.0002	0.0002
	(0.0043)	(0.0036)	(0.0004)	(0.0003)	(0.0115)	(0.0116)	(0.0010)	(0.0010)
DD30	-0.3819*	-0.2289	-0.0264*	-0.0435**	0.3334	0.5687	0.0209	0.0024
	(0.2270)	(0.1950)	(0.0146)	(0.0168)	(0.3802)	(0.3812)	(0.0214)	(0.0203)
DD1030	0.1958	0.1044	0.0228**	0.0203	0.2262	0.1845	0.0157*	0.0208*
	(0.1679)	(0.1461)	(0.0111)	(0.0135)	(0.1390)	(0.1391)	(0.0094)	(0.0113)
N	23742	23742	52101	52101	7443	7443	17307	17307
R^2			0.013	0.017			0.016	0.023
Time trend	Year	Country x Year	Year	Country x Year	Year	Country x Year	Year	Country x Year
Model	Logit	Logit	LPM	LPM	Logit	Logit	LPM	LPM
Precip. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Agri.	Agri.	Agri.	Agri.	Non-agri.	Non-agri.	Non-agri.	Non-agri.

The dependent variable is binary and captures household-level out-migration incidence in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The wealth index lies between 0 and 1, where higher values represent higher wealth. The migration and wealth variables are constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. Models 1-4 capture agricultural and models 5-8 non-agricultural households. All models further control for growing season precipitation and their squared terms. Models 1-2 and 5-6 are estimated using a Logit model and models 3-4 and 7-8 using LPM. Models 1, 3, 5 and 7 use a common and models 2, 4, 6 and 8 country-specific time trend. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

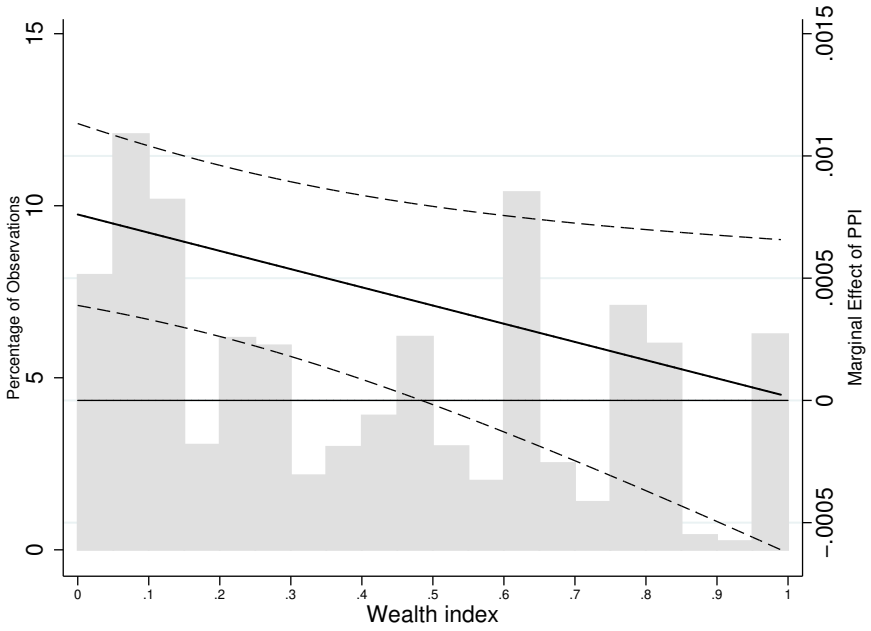


Figure 4.2: Partial effects of PPI by wealth according to model 3, Table 4.7 with 90% CIs

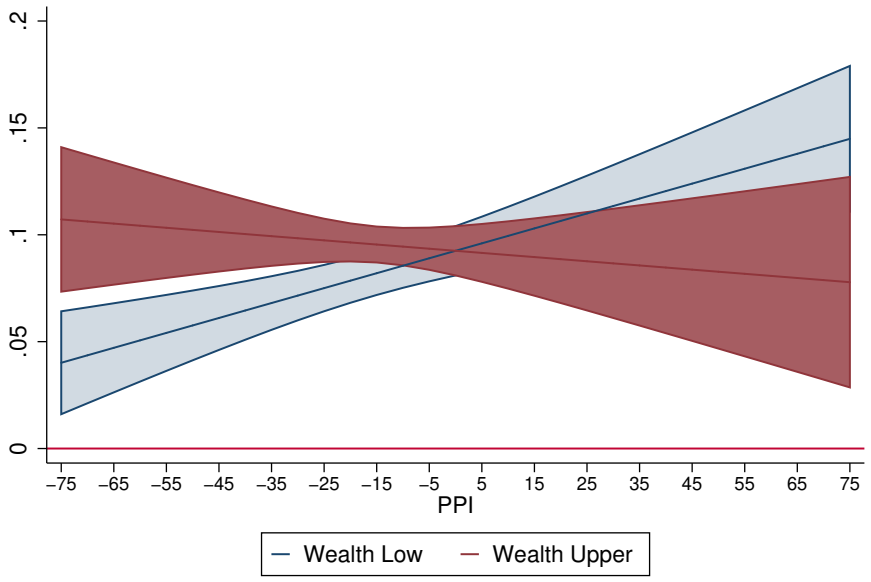


Figure 4.3: Partial effects of PPI on agricultural households by wealth according to model 5, Table 4.14 (Appendix 4.D) with 90% CIs

4.5.3 Destination choices

Throughout this paper, we have made the case for both global prices and local weather being a) exogenously determined and b) the two factors being uncorrelated. A third, albeit less crucial source of potential bias in our estimates is the spatial correlation in

production pattern and local weather conditions. Due to the necessary arbitrariness in drawing district borders, neighbouring districts tend to grow similar crops and may therefore experience income shocks that are spatially correlated. Similarly, neighbouring regions may experience climatic and weather conditions that are not entirely dissimilar across districts. We illustrate the presence of spatial correlations in our setting in figures 4.7, 4.4 and 4.5 of appendices 4.B and 4.A.

The consequence of these spatial correlations follows immediately from our characterisation of the household out-migration probability in equation 4.4: For example, spatial correlations in local production patterns mean that a positive income shock through a rise in locally-relevant global crop prices that increases the utility of the potential migrant staying home, U_{hhy}, simultaneously changes the households' expected utility flows in destination districts, U_{hdy}. Thus, the attractiveness of some nearby destinations may co-move when locally relevant global prices change. It follows that, due to spatial correlations, locally-relevant global prices can be expected to have less of an impact on internal rural migration, a type of migration that makes only for a very small share of total migration (see Table 4.1).

However, destination choices may not only be influenced by spatial correlations. Unless local production and local consumption of agricultural products are uncorrelated, the real income in urban areas will be negatively affected by increases in locally relevant global crop prices (see equation 4.7). For potential migrants in agricultural rural households, this means that with rising income at home, urban destinations in the same district (or nearby districts if spatial correlations are considered) may become less attractive whenever income at home rises through increases in producer prices. We note that the same reasoning does not hold for local weather shocks: Local weather conditions, which act on the locally produced quantity have no effect on the attractiveness of urban areas (see equation 4.8), where real income remains unaffected. In sum, these points lead us to conclude that the aggregate positive effect of locally-relevant global prices on out-migration in agricultural households is unlikely to be driven by an increase in rural-urban migration.

We test these consideration within our baseline specification (4.10) for each of the destination choices we observe in the data.[10] Table 4.8 shows the results of these analyses derived from our preferred specification. For completeness, we present the results for non-

[10]We consider each destination separately at a time and set all other destinations to zero in these analyses. An alternative econometric approach would be to condition estimates on locally-relevant global producer prices and weather conditions in nearby district, potentially weighing these by distance. However, due to the high spatial correlation of adjacent districts in our setting, the collinearity introduced by such an approach does not allow for reliable inference.

agricultural households in Appendix 4.E, Table 4.16. We also present robustness tests of specifications that include country-by-year fixed effects in the Appendix 4.E in Tables 4.17 and 4.18 gor agriculutral and non-agricultural households, respectively. Including these more demanding fixed effects does not alter our results.

	(1) Total	(2) Internal	(3) Internal: rural	(4) Internal: urban	(5) Other African	(6) OECD
PPI	0.0006***	-0.0001	-0.0000	-0.0001	0.0006***	-0.0000
	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0000)
DD1030	0.0219*	0.0233***	0.0006	0.0234***	0.0033	0.0031
	(0.0112)	(0.0081)	(0.0033)	(0.0079)	(0.0041)	(0.0048)
DD30	-0.0254*	-0.0257**	0.0002	-0.0261**	-0.0068	-0.0043*
	(0.0148)	(0.0118)	(0.0040)	(0.0100)	(0.0061)	(0.0026)
N	52101	52101	52101	52101	52101	52101
R^2	0.013	0.014	0.004	0.011	0.008	0.000
Time trend	Year	Year	Year	Year	Year	Year
Model	LPM	LPM	LPM	LPM	LPM	LPM

The dependent variables are binary and capture household-level out-migration incidence by destination in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. All models capture agricultural households, control for growing season precipitation and their squared terms and are estimated with LPM. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.8: The effect of PPI by destination choice: Agricultural households

The results confirm that the aggregate effect of changes in locally-relevant prices on household out-migration among agricultural households is driven by migration to destinations in other African countries. On the contrary, changes in local weather mostly affects internal migration into urban areas within agricultural households. While our data does not allow us to precisely pin down the effect of spatial correlations, our results are highly suggestive of a simultaneous decline in real income in urban areas when locally-relevant prices rise. Thus, while the aggregate income effect of locally-relevant global prices and local weather is positive in our sample, the type of migration induced through the income channel differs: A rise in global prices increases migration into neighbouring African countries, while an improvement in local weather conditions increases internal rural-urban migration.

It should be noted that the income fluctuations agricultural households experience from exogenous changes in global prices and local weather conditions are unlikely to be sufficient to cover the necessary investment cost of migration into OECD countries. These have been shown to be significantly higher than the costs of internal migration and migration into nearby African destinations [Marchal and Naiditch, 2020]. Dao et al. [2018] further show that for international migration flows to OECD countries, slow-moving macroeconomic drivers are more important to explain the positive association between income and emigration at early stages of development. Our results indeed suggest that

migration to these destinations are unaffected by short-term fluctuations in both prices and weather.

4.6 The role of conflict

Here, we explore whether in addition to wealth, conflict is a concurring mechanism behind the PPI - migration association, revealed in section 4.5.1. Socio-political conditions importantly affect the relationship between climate-related events and human migration [Black et al., 2011, Cattaneo et al., 2019]. Ample evidence suggests a general link between local climatic pressure and conflict [Abel et al., 2019, Hsiang et al., 2013], which can spill over into migration and displacement [Missirian and Schlenker, 2017, Abel et al., 2019]. Fluctuations in global food prices specifically have been shown to affect political instability and violent conflicts [Bazzi and Blattman, 2014, Bellemare, 2015, McGuirk and Burke, 2020, De Winne and Peersman, 2019]. Therefore, while not making general claims about wider displacement, we study if these changes in international food prices can indirectly influence the specific type of labour migration at hand, household out-migration.

As detailed in section 4.4, analysing the channeling effect of conflict in such settings is connected to numerous empirical challenges. We therefore first estimate equation 4.12 to study the general correlation between household out-migration and conflict in the setting of Sub-Saharan Africa. The results are shown in Table 4.9. Models 1-2 and 5-6 show the effects of output conflict on migration from agricultural and non-agricultural households, respectively. Models 3-4 and 7-8 then show the effects of factor conflict for agricultural and non-agricultural households, respectively. For agricultural households, we do not find any statistically significant association between output conflict and migration or between factor conflicts and migration. For non-agricultural households, we find weak evidence that household out-migration may be negatively correlated with factor conflicts. One explanation of this negative correlation could be that out-migration serves as an *escape valve* for local tensions [Bosetti et al., 2020]; however, we abstain from a strong interpretation of the result.

To complete our analyses, we also examine the general link between global commodity price changes and their effect on output and factor conflicts in Table 4.19 of Appendix 4.F. Our results reveal that a rise in producer commodity prices decreases the likelihood of output conflicts (for a more in-depth discussion, see Appendix 4.F). Thus, we find no evidence for a link between rising producer prices and conflict onset as a mechanism for

migration in our particular setting. Taken together, the association between prices and conflict previous studies have documented does not seem to play a role for the household decision to send a member as a labour migrant, the predominant type of migration in Sub-Saharan Africa. If anything, household out-migration shows a negative correlation with conflict, possibly since migration decreases the local tensions and competition over resources, as documented by previous literature [Bosetti et al., 2020].

Table 4.9: Effect of conflict on migration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Output conflict	0.0074	0.0076			-0.0002	-0.0009		
	(0.0054)	(0.0051)			(0.0052)	(0.0056)		
Factor conflict			-0.0167	-0.0118			-0.0169	-0.0160*
			(0.0127)	(0.0086)			(0.0118)	(0.0081)
N	52101	52101	52101	52101	17307	17307	17307	17307
R^2	0.012	0.016	0.012	0.016	0.016	0.023	0.016	0.023
Time trend	Year	Country x Year	Year	Country x Year	Year	Country x Year	Year	Country x Year
Model	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM
Precip. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Agri.	Agri.	Agri.	Agri.	Non.-agri.	Non.-agri.	Non.-agri.	Non.-agri.

The dependent variable is binary and captures household-level out-migration incidence in a given year. The *Output conflict* variable is binary and measures smaller scale conflict incidence. The *Factor conflict* variable is binary and measures smaller larger conflict incidence. The migration and wealth variables are constructed using World Bank's African Migration and Remittances Surveys data. *Output conflict* is constructed using ACLED data. *Factor conflict* is constructed using UCDP data. All models are estimated using LPM. Models 1-4 capture agricultural and models 5-8 non-agricultural households. Models 1, 3, 5 and 7 use a common and models 2, 4, 6 and 8 country-specific time trend. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

4.7 Concluding remarks

In this paper, we conduct household-level analyses on the relation between global prices and the household decision to send a migrant and derive a series of new conclusions on climate-related migration. We study how international crop price changes, to a large extent induced by distant climatic shocks in agricultural production, affect migration in Sub-Saharan Africa. By acknowledging the importance of transmission of climatic shocks in an interconnected world, our study provides a new perspective on climate-related migration dynamics.

We further provide new evidence on the income elasticity of migration that can affect households' ability to use migration as an adaptation strategy. We find that higher food prices can help agricultural households overcome their budget constraint. Higher producer prices thus facilitate migration similar to positive income shocks brought about by spells of good local weather conditions during the growing season. This aggregate positive effect of global prices on household out-migration is driven by the low average household wealth level in agricultural Sub-Saharan Africa. In the future, with household wealth rising over time, the opportunity cost channel is likely to take over, at least in partial equilibrium.

Our findings further show that, unlike positive weather shocks which mostly facilitate internal rural-urban migration, positive income shocks through rising producer prices only increase migration to neighbouring African countries, likely due to the simultaneous decrease in real income in nearby urban areas. This finding has important implications for projected migration dynamics in the poor regions of Sub-Saharan Africa: While we confirm that climate change scenarios rightly assume that local climatic conditions mostly drive internal migration, interconnected global shocks with repercussions for wider geographical areas may trigger migration into regions outside the home country.

The implications of our results will become increasingly important in the future and have direct policy implications. Episodes of rising food prices are expected more frequently as a result of the adverse climate change impacts on agricultural productivity [Lobell et al., 2011, Burke and Emerick, 2016]. On the one hand, higher food prices could help poorer agricultural households in particular to cover migration-related costs. Migration is often used as an important livelihood strategy for these households and a rise in income could ease existing budget constraints. On the other hand, higher international prices are likely to have adverse income effects on net consumers. These redistributive effects should be comprehensively considered by policy makers when designing policies to minimise welfare

losses in a changing climate.

One limitation applies to our work and needs to be discussed in this context. Based on our empirical results, we can only derive weak conclusions regarding the implications of rising producer prices for non-agricultural households, which mostly reside in urban areas. Data limitations do not allow the approximation of consumer prices relevant for local consumption at geographical levels similar to the district level analyses we conduct for producer prices. McGuirk and Burke [2020] try to overcome this issue by combining annual variation in global crop prices with country-level variation of consumption patterns. The authors thus calculate international consumer prices at the country-year level. In our sample of five countries and nine years of observation, this approach would leave us with 45 data points, insufficient to capture meaningful effects. Future scientific efforts should therefore aim to improve the evidence base on migration implications of international price changes for non-agricultural households who suffer losses in real income when crop prices rise.

A second limitation of this study pertains to the global prices - conflict - migration nexus. Considering migration of whole households rather than household out-migration could shed more light on how global prices may result in forced displacement. Such analysis would enable policymakers to evaluate the overall welfare effects of international price variations in a changing climate.

Appendix

4.A Local weather

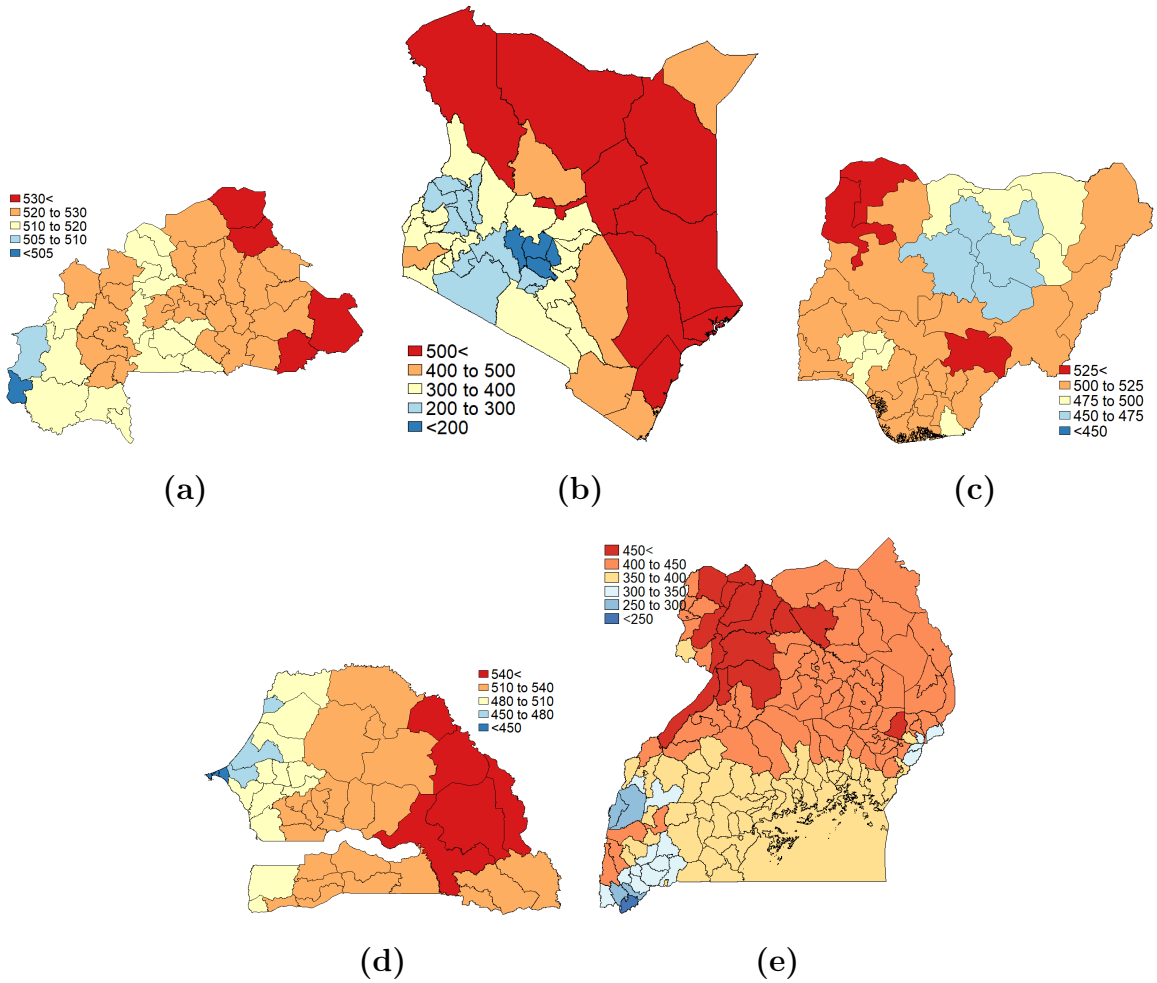


Figure 4.4: District-specific growing degree days (10-30°C): average for 2000-2008 (**(a)** Burkina Faso **(b)** Kenya **(c)** Nigeria **(d)** Senegal **(e)** Uganda)

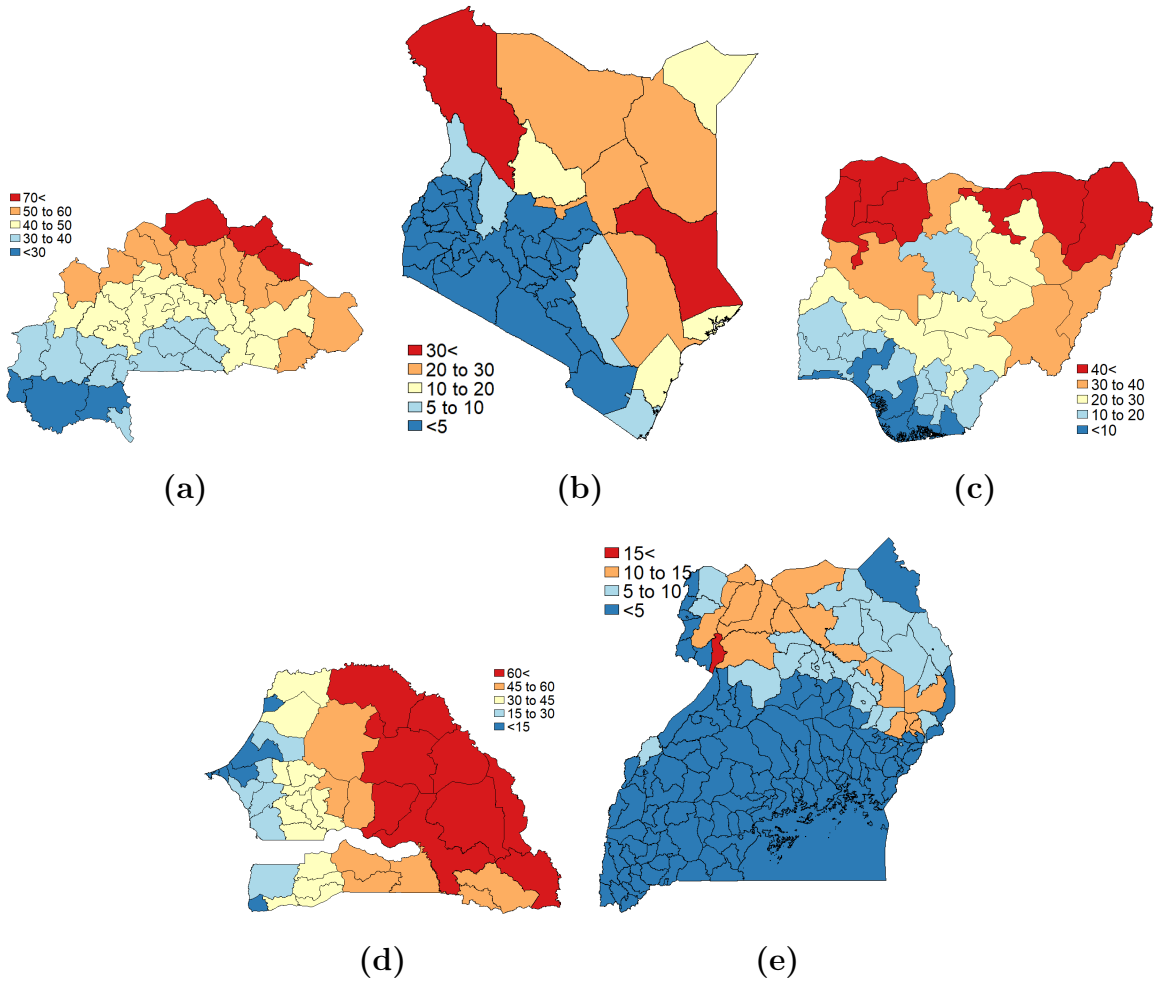


Figure 4.5: District-specific growing degree days (30°C): Average for 2000-2008 (**(a)** Burkina Faso **(b)** Kenya **(c)** Nigeria **(d)** Senegal **(e)** Uganda)

4.B International food price index

Crop	Source	PPI
Cereals		
Maize	IMF	Yes (food)
Rice	IMF	Yes (food)
Wheat	IMF	Yes (food)
Fruits and vegetables		
Soybean	IMF	Yes (cash)
Tomatoes	IMF	Yes (cash)
Vegetable oils		
Palm oil	IMF	Yes (food)
Sugar		
Raw equivalent	IMF	Yes (food)
Refined	IMF	Yes (food)
Beverages & others		
Cocoa	IMF	Yes (cash)
Coffee	IMF	Yes (cash)
Tea	IMF	Yes (cash)
Tobacco	World Bank	Yes (cash)

Table 4.10: Commodity price data to generate PPI

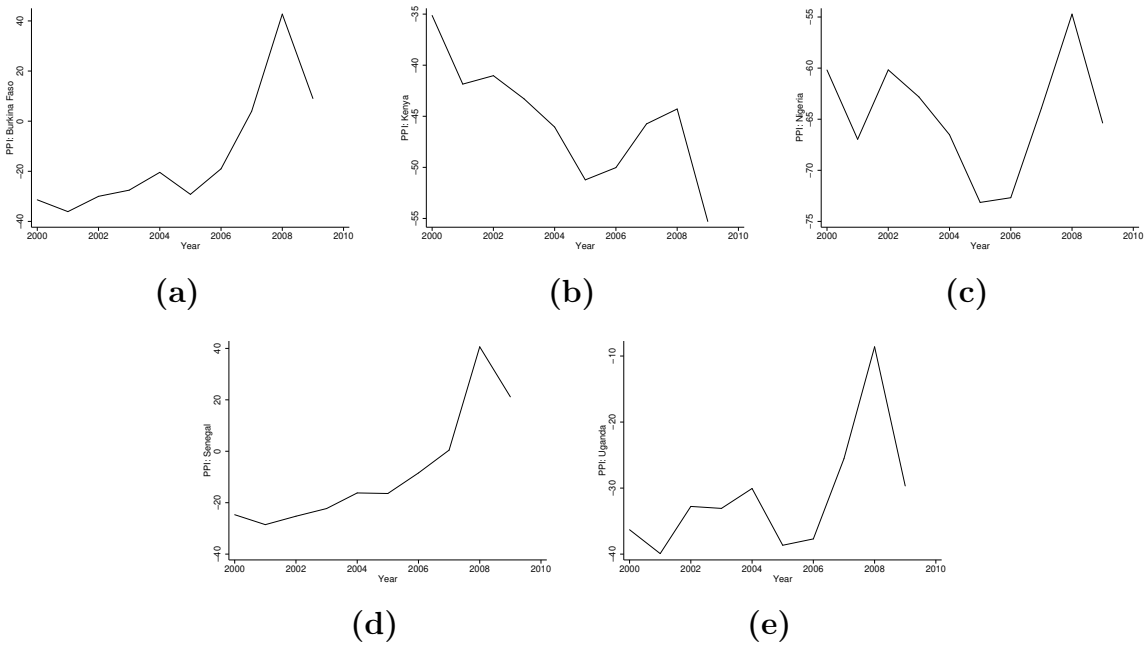


Figure 4.6: Country-specific PPI by year ((**a**) Burkina Faso (**b**) Kenya (**c**) Nigeria (**d**) Senegal (**e**) Uganda)

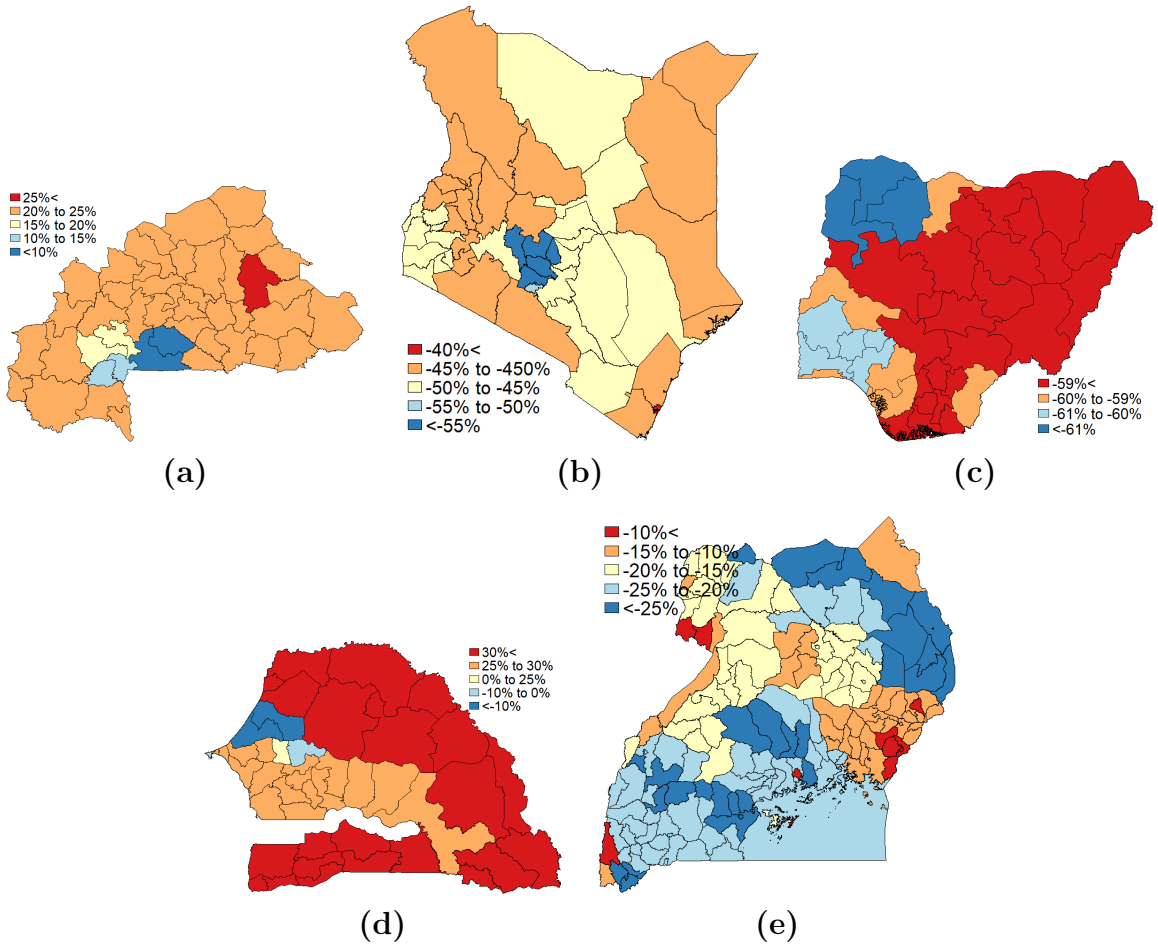


Figure 4.7: Average PPI (district-specific percentage change from the long-run average) during the 2007/08 food price crisis (**(a)** Burkina Faso **(b)** Kenya **(c)** Nigeria **(d)** Senegal **(e)** Uganda)

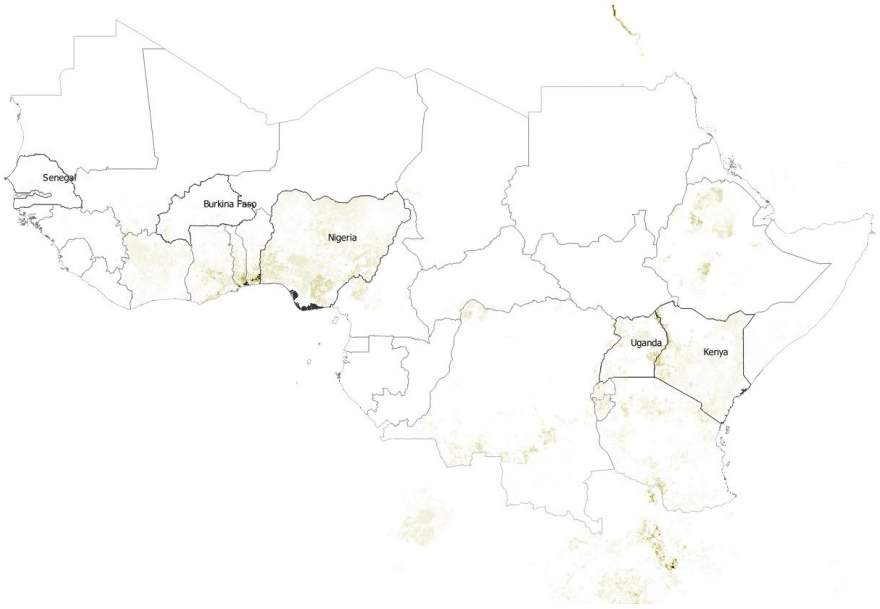


Figure 4.8: Fraction of harvested area of maize at the grid-cell level, darker color indicates higher fraction

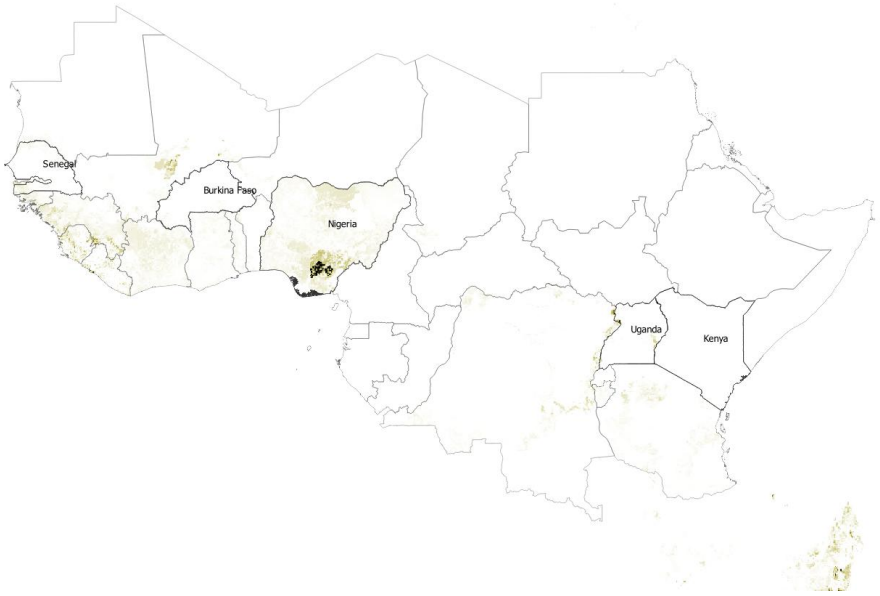


Figure 4.9: Fraction of harvested area of rice at the grid-cell level, darker color indicates higher fraction

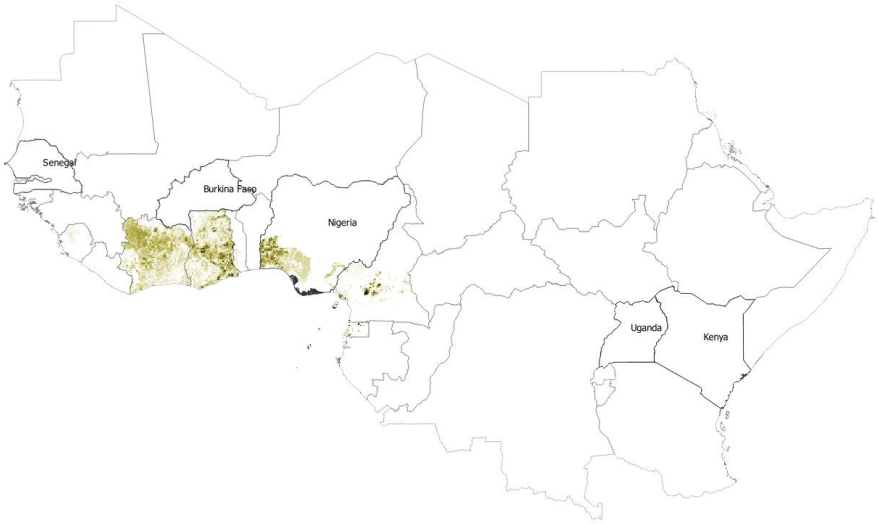


Figure 4.10: Fraction of harvested area of cocoa at the grid-cell level, darker color indicates higher fraction

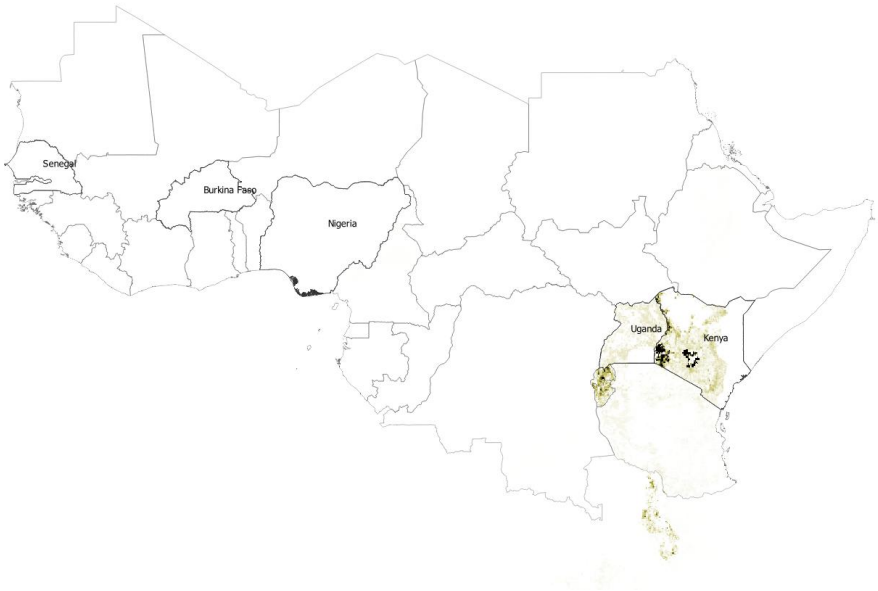


Figure 4.11: Fraction of harvested area of tea at the grid-cell level, darker color indicates higher fraction

4.C Sensitivity tests: Aggregate effects

As suggested in the main text, we further run the fully specified LPM model by distinguishing the PPI of cash and food crops. The outcomes are presented in Table 4.11. For the **sub-sample of agricultural households**, we find that the estimated effects in Table 4.5 are mainly driven by price changes of food crops. Thus, in our sample covering the food price crisis, commodities that were both locally produced and consumed were more prominently to help producing households relax their budget constraints than cash crops. For the **sub-sample of non-agricultural households**, we find weak evidence that an increase in PPI of food crops reduces the probability that households send out a migrant. PPI of food crops better captures prices of locally consumed goods, suggesting that higher food prices impose a stricter budgetary constraint on net consumers, reducing their ability to move.

Table 4.11: Effect of PPI of cash and food crops on the probability of migration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PPI (food)	0.0007***	0.0006**			0.0003	-0.0007*		
	(0.0002)	(0.0003)			(0.0003)	(0.0004)		
PPI (cash)			0.0001	0.0000			-0.0001	0.0007
			(0.0001)	(0.0003)			(0.0001)	(0.0007)
DD30	-0.0255*	-0.0460**	-0.0258	-0.0510***	0.0263	-0.0092	0.0358	-0.0039
	(0.0147)	(0.0178)	(0.0176)	(0.0193)	(0.0241)	(0.0265)	(0.0234)	(0.0265)
DD1030	0.0269**	0.0249*	0.0212*	0.0237*	0.0166	0.0247	0.0121	0.0258
	(0.0118)	(0.0142)	(0.0120)	(0.0141)	(0.0143)	(0.0159)	(0.0145)	(0.0159)
N	51075	51075	51075	51075	14427	14427	14427	14427
R^2	0.013	0.017	0.012	0.017	0.018	0.027	0.018	0.027
Time trend	Year	Country x Year	Year	Country x Year	Year	Country x Year	Year	Country x Year
Model	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM
Precip. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Agri.	Agri.	Agri.	Agri.	Non-agri.	Non-agri.	Non-agri.	Non-agri.

The dependent variable is binary and captures household-level out-migration incidence in a given year. The producer price indexes (*PPI (food)* and *PPI (cash)*) are measured in percent and capture *PPI* change (%) compared to the long-run average (1990-1999) for food and cash crops separately. *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. Models 1-4 capture agricultural and models 5-8 non-agricultural households. All models further control for growing season precipitation and their squared terms and are estimated with LPM. Models 1, 3, 5 and 7 use a common and models 2, 4, 6 and 8 country-specific time trend. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

We then turn to robustness checks on the alternative definition of the growing season. In Tables 4.12 and 4.13 we present the outcomes for agricultural and non-agricultural households respectively. First, in models 1-2 we extend the definition of the growing season by one month in each direction. Second, in models 3-4 we look at the effects of annual weather conditions. As for the agricultural households, in models 1-3 the effect of 10 to 30°C degree days remains positive, but is only significant in model 1. In model 4 the effect stays insignificant but swaps the sign. The effect of degree days above 30°C is insignificant throughout the specifications and except of model 3 it remains negative. As for the non-agricultural households, the effect of 10 to 30°C degree days remains positive but insignificant throughout the specifications. These outcomes suggest that even though the direction of the local weather effects remain mostly unchanged when using alternative growing season definitions, migration reacts significantly particularly to weather conditions during the growing season as defined in the main analysis (June-August) and suggested by broader literature. For non-agricultural households, we find new evidence that degree days above 30°C outside the growing season drive out-migration. Even though analysis of weather-related migration from non-agricultural and urban households is an existing gap in the literature [Sedova et al., 2021] this is beyond the scope of this paper and thus we abstract from the interpretation.

	(1)	(2)	(3)	(4)
PPI	0.0006***	0.0005**	0.0006***	0.0005*
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
DD30 (May-Sep.)	-0.0103	-0.0213		
	(0.0121)	(0.0130)		
DD1030 (May-Sep.)	0.0154**	0.0127		
	(0.0076)	(0.0087)		
DD30 (Annual)			0.0037	-0.0131
			(0.0067)	(0.0082)
DD1030 (Annual)			0.0033	-0.0005
			(0.0036)	(0.0049)
N	52101	52101	52101	52101
R^2	0.013	0.017	0.012	0.017
Time trend	Year	Country x Year	Year	Country x Year
Model	LPM	LPM	LPM	LPM
Precip. controls	Yes	Yes	Yes	Yes
Sample	Agri.	Agri.	Agri.	Agri.

The dependent variable is binary and captures household-level out-migration incidence in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030 (May-Sep.)* and *DD1030 (Annual)* capture 100 degree days between 10 and 30°C and *DD30 (May-Sep.) DD30 (Annual)* above 30°C between May and September and annually. The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. The sample captures agricultural households only. All models further control for precipitation and their squared terms during the considered time frame. All models are estimated with LPM. Models 1 and 3 use a common and models 2 and 4 country-specific time trend. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.12: Effects of PPI and local weather on the probability of migration of agricultural households: Alternative growing season definition

	(1)	(2)	(3)	(4)
PPI	0.0003	0.0003	0.0004	0.0003
	(0.0003)	(0.0004)	(0.0003)	(0.0004)
DD30 (May-Sep.)	0.0443***	0.0361*		
	(0.0148)	(0.0206)		
DD1030 (May-Sep.)	0.0051	0.0078		
	(0.0051)	(0.0081)		
DD30 (Annual)			0.0233**	0.0082
			(0.0117)	(0.0157)
DD1030 (Annual)			0.0038	0.0010
			(0.0034)	(0.0046)
N	17307	17307	17307	17307
R^2	0.017	0.023	0.017	0.023
Time trend	Year	Country x Year	Year	Country x Year
Model	LPM	LPM	LPM	LPM
Precip. controls	Yes	Yes	Yes	Yes
Sample	Non-agri.	Non-agri.	Non-agri.	Non-agri.

The dependent variable is binary and captures household-level out-migration incidence in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030 (May-Sep.)* and *DD1030 (Annual)* capture 100 degree days between 10 and 30°C and *DD30 (May-Sep.) DD30 (Annual)* above 30°C between May and September and annually. The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. The sample captures non-agricultural households only. All models further control for precipitation and their squared terms during the considered time frame. All models are estimated with LPM. Models 1 and 3 use a common and models 2 and 4 country-specific time trend. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.13: Effects of PPI and local weather on the probability of migration of non-agricultural households: Alternative growing season definition

4.D Sensitivity tests: Household wealth

	(1)	(2)	(3)	(4)	(5)	(6)
PPI	0.0057**	0.0059**	0.0110***	0.0007***	0.0007***	0.0005**
	(0.0027)	(0.0025)	(0.0025)	(0.0002)	(0.0002)	(0.0002)
PPI × Medium-Wealth	-0.0026	-0.0026	-0.0025	-0.0003*	-0.0003*	-0.0001
	(0.0016)	(0.0016)	(0.0017)	(0.0002)	(0.0002)	(0.0002)
PPI × Upper-Wealth	-0.0089**	-0.0088**	-0.0079**	-0.0009**	-0.0009**	-0.0004
	(0.0044)	(0.0044)	(0.0040)	(0.0003)	(0.0003)	(0.0003)
DD30		-0.3586	-0.2100		-0.0254*	-0.0438***
		(0.2265)	(0.1974)		(0.0145)	(0.0167)
DD1030		0.1783	0.0934		0.0215*	0.0204
		(0.1701)	(0.1479)		(0.0111)	(0.0135)
N	23742	23742	23742	52101	52101	52101
R^2				0.013	0.013	0.017
Time trend	Year	Year	Country x Year	Year	Year	Country x Year
Model	Logit	Logit	Logit	LPM	LPM	LPM
Precip. controls	No	Yes	Yes	No	Yes	Yes
Sample	Agri.	Agri.	Agri.	Agri.	Agri.	Agri.

The dependent variable is binary and captures household-level out-migration incidence in a given year. The wealth variable is categorical and takes on values for the year 2000: low wealth (0), medium wealth (1) upper wealth (2). The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration and wealth variables are constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. The sample captures agricultural households only. Models 2-3 and 5-6 further control for growing season precipitation and their squared terms. Models 1-3 are estimated with fixed effects logit model and models 4-6 with LPM. Models 1-2 and 4-5 use a common and models 3 and 6 country-specific time trend. Model 5 corresponds to the preferred specification. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.14: Heterogeneous effects of PPI on migration by household wealth: Agricultural households

	(1)	(2)	(3)	(4)	(5)	(6)
PPI	0.0061	0.0059	0.0102**	0.0007	0.0007	0.0006
	(0.0043)	(0.0044)	(0.0049)	(0.0005)	(0.0005)	(0.0005)
PPI × Medium-Wealth	-0.0037	-0.0038	-0.0049	-0.0005	-0.0005	-0.0004
	(0.0034)	(0.0035)	(0.0036)	(0.0004)	(0.0004)	(0.0004)
PPI × Upper-Wealth	-0.0053	-0.0053	-0.0051	-0.0007	-0.0007	-0.0005
	(0.0078)	(0.0079)	(0.0081)	(0.0007)	(0.0007)	(0.0007)
DD30		0.3779	0.6080		0.0220	0.0042
		(0.3685)	(0.3776)		(0.0208)	(0.0196)
DD1030		0.2256	0.1854		0.0157*	0.0205*
		(0.1381)	(0.1379)		(0.0094)	(0.0112)
N	7443	7443	7443	17307	17307	17307
R^2				0.016	0.017	0.023
Time trend	Year	Year	Country x Year	Year	Year	Country x Year
Model	Logit	Logit	Logit	LPM	LPM	LPM
Precip. controls	No	Yes	Yes	No	Yes	Yes
Sample	Non-agri.	Non-agri.	Non-agri.	Non-agri.	Non-agri.	Non-agri.

The dependent variable is binary and captures household-level out-migration incidence in a given year. The wealth variable is categorical and takes on values for the year 2000: low wealth (0), medium wealth (1) upper wealth (2). The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration and wealth variables are constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. The sample captures non-agricultural households only. Models 2-3 and 5-6 further control for growing season precipitation and their squared terms. Models 1-3 are estimated with fixed effects logit model and models 4-6 with LPM. Models 1-2 and 4-5 use a common and models 3 and 6 country-specific time trend. Model 5 corresponds to the preferred specification. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.15: Heterogeneous effects of PPI on migration by household wealth: Non-agricultural households

4.E Sensitivity tests: Destination choices

	(1) Total	(2) Internal	(3) Internal: rural	(4) Internal: urban	(5) Other African	(6) OECD
PPI	0.0004	0.0002	0.0000	0.0002	0.0000	0.0001
	(0.0003)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
DD1030	0.0156*	0.0165*	0.0059**	0.0106	0.0051	0.0050
	(0.0093)	(0.0094)	(0.0026)	(0.0095)	(0.0032)	(0.0071)
DD30	0.0205	-0.0186	-0.0082	-0.0106	0.0255**	0.0058
	(0.0214)	(0.0210)	(0.0073)	(0.0206)	(0.0103)	(0.0091)
N	17307	17307	17307	17307	17307	17307
R^2	0.016	0.023	0.006	0.018	0.003	0.002
Time trend	Year	Year	Year	Year	Year	Year
Model	LPM	LPM	LPM	LPM	LPM	LPM

The dependent variables are binary and capture household-level out-migration incidence by destination in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. All models capture non-agricultural households, control for growing season precipitation and their squared terms and are estimated with LPM. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.16: The effect of PPI by destination choice: Non-agricultural households

	(1) Total	(2) Internal	(3) Internal: rural	(4) Internal: urban	(5) Other African	(6) OECD
PPI	0.0004*	0.0001	0.0000	0.0000	0.0005***	-0.0000
	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
GDD1030	0.0198	0.0223**	0.0034	0.0189*	-0.0011	0.0038
	(0.0150)	(0.0102)	(0.0039)	(0.0102)	(0.0065)	(0.0060)
GDD30	-0.0473***	-0.0265**	0.0033	-0.0293***	-0.0166**	-0.0054
	(0.0167)	(0.0132)	(0.0062)	(0.0110)	(0.0070)	(0.0046)
N	54099	54099	54099	54099	54099	54099
R^2	0.017	0.018	0.005	0.014	0.009	0.002
Time trend	CountryXYear	CountryXYear	CountryXYear	CountryXYear	CountryXYear	CountryXYear
Model	LPM	LPM	LPM	LPM	LPM	LPM
Precip. controls	Yes	Yes	Yes	Yes	Yes	Yes

The dependent variables are binary and capture household-level out-migration incidence by destination in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. All models capture agricultural households, control for growing season precipitation and their squared terms and are estimated with LPM employing country-specific trends. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.17: PPI and destination choices: Agricultural households

	(1) Total	(2) Internal	(3) Internal:rural	(4) Internal:urban	(5) Other African	(6) OECD
PPI	0.0003	0.0003	0.0001***	0.0002	-0.0001	0.0001
	(0.0004)	(0.0003)	(0.0000)	(0.0003)	(0.0001)	(0.0002)
GDD1030	0.0186	0.0169	0.0106**	0.0073	0.0074	-0.0041
	(0.0121)	(0.0125)	(0.0047)	(0.0124)	(0.0049)	(0.0102)
GDD30	0.0142	-0.0072	-0.0124*	0.0039	0.0233**	0.0018
	(0.0244)	(0.0228)	(0.0073)	(0.0242)	(0.0113)	(0.0104)
N	18342	18342	18342	18342	18342	18342
R^2	0.024	0.031	0.009	0.026	0.006	0.005
Time trend	CountryXYear	CountryXYear	CountryXYear	CountryXYear	CountryXYear	CountryXYear
Model	LPM	LPM	LPM	LPM	LPM	LPM
Precip. controls	Yes	Yes	Yes	Yes	Yes	Yes

The dependent variables are binary and capture household-level out-migration incidence by destination in a given year. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). The migration variable is constructed using World Bank's African Migration and Remittances Surveys data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. All models capture non-agricultural households, control for growing season precipitation and their squared terms and are estimated with LPM employing country-specific trends. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Table 4.18: PPI and destination choices: Non-agricultural households

4.F Further tests on the link between producer prices and conflict

In this section, we present outcomes from a set of LPM regressions with and without state-specific trends, analysing the association between PPI, local weather and the probability of conflict incidence. We conduct the analyses at the district level. In models 1-2 and 5-6 of Table 4.19, we study the direct effects on output and factor conflict likelihoods, respectively. In models 3-4 and 7-8 of Table 4.19 we then interact the PPI and local weather variables with a district-specific fraction of agricultural households, to study if the effect of international prices varies by districts' dependence on agricultural production.

From the theoretical perspective, the direct effect of higher food prices on conflict is a priori not clear. On the one hand, the predation (or rapacity) and deprivation theories imply that higher food prices can result in more violent events. The so-called predation effect suggests that higher prices increase the value of the appropriable surplus, leading to more conflicts [Besley and Persson, 2008, Dube and Vargas, 2013]. The deprivation effect indicates that among consumers, an increase in prices can induce perceptions of relative deprivation in comparison to others and thus lead to public unrests [Hendrix and Haggard, 2015]. On the other hand, the opportunity costs effect suggests that higher food prices reduce conflicts, by increasing the opportunity costs of insurrection for farmers as higher wages and revenues make it more attractive to work [Bazzi and Blattman, 2014, Dube and Vargas, 2013, De Winne and Peersman, 2019]. Moreover, higher commodity

prices can increase state revenues and so the capacity of the state to reduce conflicts [Besley and Persson, 2008, De Winne and Peersman, 2019].

Our results reveal that a rise in producer commodity prices decreases the likelihood of output conflicts. This is in line with findings by Brückner and Ciccone [2010] or Berman and Couttenier [2015] and the opportunity cost theory [Bazzi and Blattman, 2014, Dube and Vargas, 2013]. These findings contrast the evidence by Bellemare [2015], Hendrix and Haggard [2015], Raleigh et al. [2015] or De Winne and Peersman [2019] in line with the predation (also called rapacity) and deprivation effects, both of which outline how higher food prices can increase violence. The magnitudes of the interaction terms between output conflict and agricultural dependence are very close to zero.

The direct effect of yield decreasing temperatures (*DD30*) on the likelihood of output conflict is negative and becomes significant only in models 3-4. When interacted with the agricultural dependence, the effect is positive suggesting that yield decreasing temperatures are more likely to increase the probability of output conflict in areas that are more dependent on agricultural production. This outcome is in line with the findings on PPI, namely that the opportunity costs of violence decrease with decreasing agricultural incomes [Koubi, 2019].

The effect of yield enhancing temperatures (*DD1030*) on output conflict likelihood is positive in models 1 and 3 but looses its significance in model 2 and additionally swaps the sign in model 4. When it comes to factor conflict, the effect of yield enhancing temperatures is positive throughout specifications but it only becomes significant in model 5. These generally positive associations are in line with the predation theory and seem to be relevant primarily for net-consumers (the interaction with agricultural dependence is insignificant), for whom the increase in agricultural surplus increases the rewards from engaging in conflict.

Table 4.19: Effect of PPI on conflict

	Output conflict				*Factor conflict*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PPI	-0.001233*	-0.000677	-0.002686*	-0.003222*	-0.000408	0.000377	0.000012	0.000603
	(0.000689)	(0.000967)	(0.001444)	(0.001782)	(0.000485)	(0.000638)	(0.000818)	(0.000926)
DD1030	0.131833***	0.005112	0.156017*	-0.079531	0.076543**	0.018637	0.077876	0.078093
	(0.049673)	(0.066760)	(0.087221)	(0.095586)	(0.038051)	(0.048522)	(0.091138)	(0.090370)
DD30	-0.092142	-0.106884	-1.235795***	-1.429446**	-0.112580	-0.102938	-0.354520	-0.398344
	(0.075057)	(0.143670)	(0.467423)	(0.591827)	(0.069739)	(0.088839)	(0.302292)	(0.259986)
PPI × Agricultural (%)			0.000021	0.000032*			-0.000005	-0.000003
			(0.000014)	(0.000017)			(0.000007)	(0.000007)
DD1030 × Agricultural (%)			-0.000295	0.001175			-0.000031	-0.000755
			(0.001240)	(0.001361)			(0.001167)	(0.001126)
DD30 × Agricultural (%)			0.012997***	0.014654**			0.002724	0.003311
			(0.004928)	(0.005820)			(0.003243)	(0.002746)
N	1260	1260	1260	1260	1260	1260	1260	1260
R^2	0.045	0.125	0.052	0.134	0.011	0.043	0.012	0.044
Time trend	Year	Country x Year	Year	Country x Year	Year	Country x Year	Year	Country x Year
Model	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM
Precip. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Conflicts	Jul.-Dec.	Jul.-Dec.	Jul.-Dec.	Jul.-Dec.	Jul.-Dec.	Jul.-Dec.	Jul.-Dec.	Jul.-Dec.

The dependent variable is binary and captures district-level conflict incidence in a given year. The *Output conflict* variable is binary and measures smaller scale conflict incidence. The *Factor conflict* variable is binary and measures larger conflict incidence. The producer price index (*PPI*) is measured in percent and captures *PPI* change (%) compared to the long-run average (1990-1999). *DD1030* captures 100 degree days between 10 and 30°C and *DD30* above 30°C during the growing season (June-August). *Output conflict* is constructed using ACLED data. *Factor conflict* is constructed using UCDP data. Weather variables are constructed using ERA5 data. PPI is constructed by combining crop-specific fraction of harvested area data by Monfreda et al. [2008] and annual global commodity prices from the IMF International Finance Statistics series and the World Bank Global Economic Monitor. All models are estimated using LPM. Models 1, 3, 5 and 7 use a common and models 2, 4, 6 and 8 country-specific time trend. Standard errors clustered at the district level are displayed in parentheses.* p<0.10, ** p<0.05, *** p<0.01.

Chapter 5

Immigration and redistributive spending: Evidence from English local authorities

5.1 Introduction

Concerns about redistributing income to what are considered outsiders has featured as a salient issue in the run up to the 2016 Brexit referendum that ultimately saw the UK leaving the European Union (EU) by popular vote. Both UK transfers to the EU and the pressure EU immigrants allegedly put on public service provision in the UK were platforms the "Vote Leave" campaign heavily relied on to mobilise its supporters [Gherghina and O'Malley, 2019, Goodwin and Milazzo, 2017, Becker et al., 2017].

In this paper, we investigate if the link between the inflow of EU "outsiders" and a local loss of appetite for redistributing income was visible in the UK before the Brexit referendum took place. Specifically, we focus on the time period after the 2004 and 2007 EU enlargement when the historically anchored EU scepticism in the UK took a turn against migration following the large inflows of Central and Eastern Europeans (AC-12) into UK territory [Gherghina and O'Malley, 2019]. We analyse the effect of the unexpected and spatially heterogeneous AC-12 migration wave on English local authority level public spending and revenue to answer the question if the local presence of migrants is indeed associated with less redistributive spending patterns. We combine detailed local authority revenue and expenditure data from the Chartered Institute of Public Finance and Accountancy (CIPFA) with annual data on local authority level migrant stocks calculated from

the UK Labour Force Survey/Annual Population Survey (LFS/APS) obtained under a special licence agreement over the 2000 to 2015 observation period. Due to the very low number of AC-12 migrants present in England in the pre-enlargement period, the estimated coefficients we obtain from our two-way fixed effects panel regressions correspond to a difference-in-differences research design, allowing us to recover treatment-on-the-treated effects of AC-12 migration on local authority revenue and spending.

Taken at face value, we find ambiguous results with regards to the hypothesis of a closed in-group that cuts down on redistributing income when faced with outsiders. On the one hand, AC-12 migration did not affect local authorities' per capita spending. However, once we zoom in on local revenue and spending patterns, we find that the presence of AC-12 migration is associated with a decrease in locally generated revenue and the unchanged per capita spending was heavily supported by an increase in funding local authorities received from the central government. AC-12 migration is further associated with both a decrease in means-tested social care spending per capita and an increase in education expenditure per capita, an expenditure item where inter-group transfers are likely to be relatively more salient compared to other non-means-tested services [Tabellini, 2020a, Speciale, 2012].

To further explain these results, we then disentangle local preferences for less redistribution from mechanical changes in demand for local services brought about by the distinct characteristics of AC-12 migrants. Our results show that the strong association of AC-12 migration with a decline in social care and a rise in education expenditure per capita is driven by changes this migrant group caused to local demographics. In fact, social care expenditure per population aged 65 and above, the main recipient population of social care services, increased strongly in response to AC-12 inflows. Similarly, when normalising education expenditure per the rising local number of pupils in areas more strongly affected by AC-12 migration, expenditure remained vastly unaffected. Thus, the effects AC-12 migrants had on local authority expenditure patterns were in large parts due to the shifts to local demographics these migrants caused and the corresponding institutional responses that were triggered by the resulting changes in local service demand. The relative shifts from social care towards education expenditure further explain the observed dynamics on the revenue side: In England, local authority education expenditure is almost entirely funded through central government grants, while a larger share of social care spending comes from locally generated revenue [Phillips, 2018]. It thus appears that migration decreased the pressure local authorities faced on social care spending over our observation period, allowing local authorities to take better care of their vulnerable older

population while decreasing the need for raising revenue locally. On the other hand, we do not find evidence that would support the hypothesis of a shift in local preferences towards less redistribution in response to migration inflows, which we test by analysing local voting patterns and native out-migration ("voting with their feet") following the 2004 EU enlargement and the subsequent wave of AC-12 migration. Our results show that a larger presence of AC-12 migrants is associated with a rise in local Council seats held by the more redistributive Labour party (rather than the Conservatives) and a decline in net-migration outflows. We note that the latter result, coined "fiscal externality" by Tabellini [2020b], may also provide a partial explanation for the changes in local demographics associated with AC-12 migration.

England is a suitable test bed for the local level link between migration and public spending for a number of reasons. First, the country experienced large waves of migration in recent decades, including the both unprecedented and unexpected wave of AC-12 migrants following the 2004 enlargement of the European Union [Becker et al., 2016]. These successive waves have led scholars to test the impact of the intensity of migration flows to the UK on numerous outcomes including crime rates [Bell et al., 2013, Jaitman and Machin, 2013], house prices [Sá, 2015], hospital waiting times [Giuntella et al., 2018], public budgets and the wages and employment of natives [Card et al., 2012, Manacorda et al., 2012, Dustmann et al., 2013, Becker et al., 2018]. In this study, we exploit the large and spatially heterogeneous shock to local migration stocks that stemmed from EU enlargement in 2004 and led to more than one million people migrating from Central and Eastern Europe to the UK. Second, the discretion in raising revenue and spending decisions at the hands of local authorities in the UK makes it an appropriate case study. Funded through a mix of central government grants and locally raised revenue, England, and the UK more generally, is one of the European countries where local governments have discretion over spending decisions that encompass several public expenditure items. UK local authorities are responsible for policies concerning education, social care services, highways, roads and transport, housing, cultural services, environmental services, planning and development and protective services [Gavazza et al., 2019, Sandford, 2018, Phillips, 2018]. Local authorities are required to balance their budget but can increase or decrease their total spending through steering the local council tax, a property tax. They can further shift their spending between more or less redistributive spending items, with spending on means-tested social care services in particular reflecting the redistribution of income towards the relatively less wealthy. Third, due to the limited scope the UK central government had under EU law to target EU migrants directly over our 2000 to 2015

observation period, the central government could not use restrictive migration policies in the years prior to the Brexit referendum, such that cutting public spending was the only possible response to a decrease in appetite for redistributing resources. Finally, the wealth of information available allows to disentangle local preferences and fiscal externalities from a more mechanical migrant demand channel, a mechanism frequently neglected in the literature when studying the effect of migration on preferences for redistribution.

5.1.1 Contribution to the literature

Our results contribute to the literature on the impact of migration on redistribution in destination areas. Pioneered by Freeman [1986], Alesina et al. [2001], Alesina et al. [2004] and Easterly and Levine [1997], an important stream of literature argues that redistributive policies are supported more strongly by homogeneous groups. These findings are driven by in-group biases which translate into greater immigrant diversity lowering preferences for redistribution. Following this work, several scholars have analysed the relationship between immigration and preferences for redistribution in the context of migration to the US and European countries with results pointing towards a negative association of the two [Senik et al., 2009, Hopkins, 2010, Speciale, 2012, Halla et al., 2017, Dinas et al., 2019, Steinmayr, 2020]. For example, Senik et al. [2009] finds some evidence of a negative association studying Europe as whole. Speciale [2012] leverages the variation in inflows of migrants to EU-15 countries stemming from the 1990s Balkan wars to study the impact of migration on education expenditure and finds a small and negative association between perceived migration and support for the welfare state. Both authors document considerable heterogeneity across countries and stress the importance of sub-national studies to understand the mechanisms at play and allow for a causal investigation.

Some work at the sub-national level exists. Dahlberg et al. [2012] exploit a refugee dispersal mechanism in Sweden and find a significant negative effect of immigration on the local support for redistribution. In Denmark, Harmon [2018] uses an instrumental variable strategy based on historical housing stock data and finds that greater migration inflows leading to increases in local ethnic diversity shift election outcomes from traditional "big government" left-wing parties towards anti-immigrant nationalist parties. Both studies suggest immigration may lower the level of redistribution or public spending but identify further examination of the effect of immigration and ethnic diversity on more direct measures of redistributive spending as an important topic for future work. Similarly, in more recent work studying the effect of extending the voting franchise to non-natives,

Ferwerda [2021] stresses that while evidence points towards European citizens preferring less redistribution with greater migration, the evidence that this leads to a reduction on public good provision is less well understood. Our work contributes directly to filling this gap by measuring the impact of a large migration shock to England on local level redistribution.

Closest to this present work are studies by Tabellini [2020b], Tabellini [2020a], Gerdes [2011] and Jofre-Monseny et al. [2016]. Tabellini [2020b] studies the first "Great Migration" when 6 million black Americans migrated from the South to the North of the US. The author specifically focuses on its impact on local public finances due to changes in racial heterogeneity in Northern US cities between 1915 and 1930. After collecting data on local finances for the years 1910, 1919, and 1930 and deploying a version of a shift-share instrument based on historical settlements similar to Card [2001] and Boustan [2010] to predict black migration, the author finds that larger inflows had negative impacts on both public spending and tax revenues. The author then investigates whether this result is driven by a change in local preferences or by second-order effects black migrants had on out-migration of white Americans. While Tabellini [2020b] acknowledges that these mechanisms are not necessarily mutually exclusive, the author argues that the study's results are rather driven by a negative fiscal externality due to white flight, corroborated by the fact that cities did not change their allocation of spending. In a second related study, Tabellini [2020a] jointly investigates the political and economic impact of European immigration to the US between 1910 and 1930, using a similar shift-share instrumental variable strategy. The author finds that reductions in redistribution stemming from greater migration inflows were more likely driven by natives' preferences and cultural distance. Our study distinguishes itself from these contributions by investigating the role of migration at a time where levels of discretion of municipal councils in Europe differ from the early 20th century US. In addition to the mechanism linking migration to local expenditure and revenue offered by Tabellini [2020b] and Tabellini [2020a], we are able to study a third potential mechanism, namely the change stemming from mechanical demand for local services induced by distinct demographic characteristics of migrants.

Exploiting a refugee placement policy in Denmark, Gerdes [2011] examines the effect of immigration on municipal redistributive spending and does not find any evidence of a change in public sector spending. However, as highlighted in Harmon [2018], the author's empirical strategy might suffer from the endogenous relocation choices of immigrants not covered by the policy as well as the political discretion in assigning migrants to different municipalities. We tackle the empirical issue of endogenous location choices

of AC-12 migrants in England by presenting parallel trends comparing heavily and less heavily affected local authority areas in our main outcomes. We further show that all our results are robust to (i) matching local authorities on a wide range of 2001 Census characteristics and (ii) deploying a shift-share instrumental variable strategy based on historical settlement of AC-12 migrants in the tradition of Card [2001].

Jofre-Monseny et al. [2016] focus on the link between municipality-level variation in extra-EU immigrant density and local social spending in Spain. Instrumenting migration flows between 1998 and 2006 using the distribution of rental housing in 1991, the authors find that per capita social spending increases less in municipalities that experience the largest increases in immigrant density. While the authors report strong first-stage results in their instrumental variable regressions, one main concern with this particular identification strategy relates to the exclusion restriction. Municipalities with a relatively large supply of rental housing six years before the migration inflows are likely to also be poorer and larger than other cities and therefore may lie on differential trends that the authors do not account for. Municipalities' public finances might be affected in a way that social services spending may slow down six years later for reasons other than migration. A second concern relates to the authors' interpretation of their results. While Jofre-Monseny et al. [2016] do not study the impact of elections on native outflows and electoral outcomes, they interpret their findings as evidence for a materialisation of a shift in preferences for redistribution among the native population. We argue that this interpretation, although in line with predictions of in-group-out-group theories, is only one possibility. A decrease in redistributive public spending could also reflect a mechanical relationship introduced by migrants' socio-economic characteristics and/or their differing propensity to take up social services if inflows are sufficiently large and migrant characteristics are distinctly different from the local population.

The remainder of this paper is structured as follows. In the following section 5.2, we introduce the institutional setting of local authority spending and describe the nature of AC-12 migration flows. Section 5.3 discusses our data sources. Section 5.4 lays out our main identification strategy. Section 5.5 presents and discusses our results. Section 5.6 discusses alternative identification strategies and robustness. Section 5.7 discusses the main mechanisms of our results and section 5.8 concludes.

5.2 Institutional setting

In this section, we first provide an overview of the level of discretion local authorities hold over spending and revenue in subsection 5.2.1. We then describe the nature of AC-12 migration to England in more detail in subsection 5.2.2.

5.2.1 Local authorities in England

In total, there are 353 local authorities in England. England's local government structure is not homogeneous across the country. Local governments either function under a two-tier or a single-tier regime. The two-tier local authorities consist of an upper-tier, the county councils, and a lower-tier, the district councils. Single-tier authorities encompass 55 unitary authority councils, 36 metropolitan boroughs, 32 London boroughs, the Common Council of the City of London and the Council of the Isles of Scilly for a total of 125. Of these, we exclude nine new unitary authorities from our analyses that were only formed out of tow-tier authorities between 2007 and 2009. Two-tier authorities consist of 27 county councils which in turn are divided into 201 district councils.

While single-tier authorities bear the responsibility for all service spending decisions, county councils and district councils divide responsibilities between themselves in two-tier authorities [Sandford, 2018]. To make local authorities comparable across England, we aggregate all lower-tier authority spending up into the upper tiers. This is unproblematic for two reasons. First, the areas where spending decisions are not clearly distinguishable only concern spending on cultural goods such as museums, transport planning, economic development and tourism. Second, spending decisions on the largest expenditure items such as education and social care - the focus of our study - are made on the upper-tier county council level. We provide more information on how areas of responsibility are divided between the two tiers in 5.B of the appendix.

On the revenue side, UK local authorities obtain their funding via a mix of specific, ring-fenced grant funding, general grant funding, the collection of business rate revenue and income from a local council tax. While there is currently a trend towards more devolution in revenue, particularly regarding the retention of local business rates, the discretion of local authorities was limited to steering the local council tax rate over our main observation period from 2000 to 2015 [Phillips, 2018]. The council tax is a property tax collected by local authorities and its amount is based on the value of property. Each

property is categorised into one of eight bands (A to H) and the tax is then due annually as a fixed fraction of local authority defined baseline tax band D. In 2001, the average band D council tax rate stood at GBP 898, rising to GBP 1,459 in 2015. Total tax income collected by local authorities is then simply the multiple of the band D tax rate and the tax base, i.e. the number of band D equivalent dwellings. Over our observation period, council tax income covers about 25% of annual total service expenditure. We summarise the other main sources of income from central government grant and centrally redistributed business rates in a 'central government transfers' measure.

The discretion of the elected Councillors over local authorities' spending is not limited to revenue collected from council tax, as only a small share of funding from central government grants is explicitly ring-fenced [Phillips, 2018]. England, and the UK more generally, is one of the European countries where local governments have discretion over spending decisions that encompass several public expenditure items. UK local authorities are responsible for policies concerning education, social care services, highways, roads and transport, housing, cultural services, environmental services, planning and development and protective services such as fire and rescue [Sandford, 2018]. In this context, it is important to note that, unless local authorities can temporarily draw on previously accumulated reserves, they are required to balance their budgets and are unable to borrow on financial markets.

In 2000, British local public spending represented GBP 113 million, a value that increased to GBP 198 million in 2015 in current prices and represents approximately 25% of total government spending.[1] English local authorities spend by far the largest share of their total service expenditure on education. In 2001, education made for 53% of all service expenditure, a value that has decreased slightly over time. The second largest share of total expenditure is spent on social care services (23.5% in 2001) which has increased over time. All other expenditure items combined make for less than one fourth of total service expenditure throughout our observation period.

In our analysis, we are particularly interested in means-tested services for which local authorities have discretion during our sample period. Social care services in particular fulfil these criteria while education partially fulfils them.

Social care services in the UK consist of adult social care and child social care. Adult social care entails a range of support services available to the physically or mentally

[1]OECD (2018a), 'Consolidated government expenditure as a percentage of total general government expenditure (consolidated), http://www.oecd.org/ctp/federalism/table5_gov_exp-tot_gov_exp.xls (accessed 14/05/2021).

impaired as well as other items where the level of uptake is most highly correlated with advanced age. It also assists disadvantaged groups such as asylum seekers or substance abusers. Adult social care accounts for the bulk of social care service expenditure in England over our main observation period (approximately 70% on average). Social care service minimum eligibility criteria are set by the central government but the amount spent on social care is at local authorities' discretion [Simpson, 2017]. Phillips [2018] notes that Councils' discretion extended to determining what kind of services were offered, needs' assessments and eligibility criteria over the 2000 to 2015 observation period. The latter included different thresholds for what Councils considered the risk for an individual in absence of treatment. A detailed overview of local authority social care services and its means-testing criteria is provided in table 5.16 of appendix 5.B.

While there has been a trend towards centralisation, the sample period we study still left room for local authority discretion when it comes to education expenditure. Education expenditure in the U.K. has traditionally been locally-managed although school funding was ring-fenced as of 2006 via the Dedicated Schools Grant (DSG), which de facto represents a minimum threshold below which school spending cannot fall for most of the sample period we study. However, local authorities could still exert upward discretion and use their own revenues to additionally fund education. Recent moves by the Conservative government to remove local governments from the education expenditure formula and only have a national funding formula, whereby schools with similar characteristics receive equal funding, are yet to come into place. As highlighted by Phillips [2018], education expenditures was still partially locally managed in our sample period.

In summary, for the purposes of this paper, two observations are therefore important: First, local authorities do have a significant amount of discretion over both the revenue they collect and the allocation of their funds across different spending areas but need to balance their budgets. Second, 1100 statutory spending requirements limit local authorities in their spending decisions to some extent and do not always allow for a clear distinction between mandatory and discretionary spending [Gray and Barford, 2018]. This means that the larger the level of disaggregation of spending items, the more detailed knowledge of statutory spending requirements is necessary. We conduct our analysis on expenditure items aggregated at a relatively high level within the different spending areas to minimize this risk.

A local fiscal response to migration in England that reflects a change in redistribution could thus become visible through two main channels. First, migration could affect total

spending per capita and revenue. Second, less redistribution could also become visible through a shift between expenditure items more strongly associated with redistribution and those less associated with redistribution. In the analysis below, we focus on social care spending per capita as the main redistributive item due to its free availability only to low-income individuals. It is less clear to which extent education spending falls under redistribution. Some authors have argued that inter-group transfers are more salient in education expenditure than in other non-means-tested services [Tabellini, 2020a, Speciale, 2012]. In our analyses, we therefore leave education as a separate expenditure item. We then aggregate all expenditure that falls outside of social care and education into a third category.

5.2.2 The AC-12 migration shock

Following the EU accession of the so-called A-8 countries consisting of the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Slovakia, Poland and Slovenia, as well as Cyprus and Malta in 2004 and Bulgaria and Romania in 2007, the UK experienced a large labour migration inflow of a migrant group commonly referred to as AC-12. Due to the fear of mass migration from the, on average, less wealthy accession countries, EU member states were given the option to temporarily restrict the fundamental freedom of movement of people originating from the 2004 Central and Eastern European joiners. This was possible "under provisions in the Accession Treaty allowing the existing Member States to apply national measures regulating access to their labour markets for nationals of A-8 countries for up to seven years" [Kennedy, 2011, p.5].

All EU member states but Sweden, Ireland and the UK applied these regulations on A-8 migrants [Anderson et al., 2006]. Becker et al. [2016, p.11] explain that the decision by Tony Blair's labour government not to limit labour market access to A-8 migrants was driven by "a thriving economy and a misunderstanding of the consequences of decisions by other big EU countries to keep their borders closed to Eastern European workers for a transition period". As highlighted by the authors, a UK Home Office commissioned study conducted in 2003 by Dustmann et al. [2003] computed their migration projections under the false assumption that other countries such as Germany would also open their borders in 2004. The government then failed to adjust their initial projections when other EU member states did not open their border immediately and estimated that "only around 5,000-13,000 Eastern Europeans would arrive to the UK per year" [Dustmann et al., 2003] (as cited in Becker et al. 2016, p.11). However, between 2004 and 2007, more than 500,000

migrants arrived to the UK from Central- and Eastern Europe, vastly exceeding the initial projections [D'Auria et al., 2008].

Thus, overnight on 1 May 2004, workers from the A-8 accession countries, Malta and Cyprus obtained full rights to live and work in the UK, including the right to stay permanently and the right to be joined by dependants [Anderson et al., 2006]. Workers from A-8 countries (but not Malta and Cyprus) were obliged to register on the so-called Workers Registration Scheme (WRT). Registration on WRT gave A-8 workers access to in-work benefits but they only had access to out-of-work benefits after they had been in registered employment for 12 months [Kennedy, 2011]. For the purposes of this study, it is important to note that local authority services, including social care and education, were accessible to AC-12 workers as soon as they were registered on WRT.

Table 5.1 summarises the AC-12 migrants' socio-economic characteristics relative to British natives, EU-15 migrants and the rest of the world based on the 2011 UK Population Census. The table shows age and employment characteristics that could potentially have had an impact on the local population's attitude towards AC-12 migrants. Both employment and young age may thereby reflect a lower likelihood of welfare dependency [Gherghina and O'Malley, 2019, Goodwin and Milazzo, 2017, Becker et al., 2017]. As noted above, these characteristics may also reflect the group's need for local authority services. The migration inflow to the UK stemming from the EU enlargement was indeed both sizeable and characterised by the distinctive features of AC-12 migrants in terms of age, education and employment.

Most of the AC-12 migrants living in the UK in 2011 were in the 25 to 39 years' old categories. This contrasts sharply with the age structure of AC-12 migrants we observe in the 2001 UK Population Census where the distribution was much flatter. Other migrants from EU-15 countries and the rest of the world are also younger than the UK born on average but their age-distribution is less skewed to the left than that of AC-12. Cross-checking these numbers with those from the 2001 Census shows little to no movement. Taken together, this suggests that the inflow of migrants from AC-12 countries was distinctively younger than AC-12 migrants living in the UK pre-enlargement and that this change in the pattern of their age-structure was distinct to this group of migrants.

In terms of qualifications, the bulk of AC-12 migrants living in England in 2011 were categorised in the "apprenticeships and other" qualifications section that does not directly translate into the UK system of qualifications but is indicative of relatively low and medium skills. It is further worth noting that the 24% share in the highest skill category

did not translate into a similar share of employment in high-skill professions [Drinkwater et al., 2009]. Most of the residents born in AC-12 countries as per the 2011 census were working in routine or semi-routine occupations, again contrasting with the 2001 situation for AC-12 migrants residing in England in 2001 and other groups in both 2001 and 2011. Thus, at least over the years following the 2004 accession, the AC-12 migration flow was in its majority a labour migration flow into low and medium skills' professions. 80% of AC-12 migrants were economically active in 2011, a share significantly above that of UK born and other migrant groups.

Age structure	UK	EU-15	AC-12	Rest of the world
Age 0 to 15	21%	12%	10%	7%
Age 16 to 24	12%	13%	15%	12%
Age 25 to 34	12%	23%	41%	23%
Age 35 to 49	20%	26%	20%	29%
Age 50 to 64	18%	13%	9%	18%
Age 65 and over	17%	13%	5%	10%
Gender				
Males	49%	46%	48%	49%
Females	51%	54%	52%	51%
Highest level of qualification				
No qualifications	23%	13%	16%	19%
Level 1 qualifications	14%	7%	7%	9%
Level 2 qualifications	16%	9%	8%	9%
Level 3 qualifications	13%	10%	7%	9%
Level 4 qualifications and above	26%	40%	24%	37%
Apprenticeships and other qualifications	7%	21%	37%	18%

Family structure	UK	EU-15	AC-12	Rest of the world
No children in family	29%	34%	35%	24%
One dependent child in family	20%	20%	29%	22%
Two or more dependent children in family	35%	34%	27%	38%
All children in family non-dependent	16%	12%	9%	16%
Economic activity				
Economically active: Total	63%	69%	80%	63%
Economically active: In employment: Total	59%	64%	76%	57%
Economically active: In employment: Full-time	43%	50%	61%	42%
Economically active: In employment: Part-time	15%	14%	15%	15%
Economically active: Unemployed: Total	5%	5%	4%	7%
Economically inactive: Total	37%	31%	20%	37%
Economically inactive: Retired	23%	15%	6%	11%
Economically inactive: Student (including full-time students)	5%	9%	5%	9%
Economically inactive: Looking after home or family	3%	4%	4%	8%
Economically inactive: Long-term sick or disabled	4%	2%	1%	3%
Economically inactive: Other	2%	2%	3%	5%

Notes: Own table based on data from the 2011 UK Census

Table 5.1: Socio-economic characteristics of immigrants vis-à-vis UK citizens and other migrants in 2011

Finally, it should be noted that AC-12 migration inflows into the UK were also geographically and compositionally different from migrants of AC-12 countries that had settled in the UK prior to the 2004 enlargement shock [Becker et al., 2018]. Before the AC-12 countries joined the EU, the stock of individuals who were born in any of the ten Central- and Eastern European accession countries was around 193,180. Unlike AC-12 migrants arriving in the UK after 2004, these migrants were mostly concentrated in the London area [Becker et al., 2016]. Approximately 30% of this group had arrived before 1981 and consisted mostly of people born in Poland, who made up 42% of the stock of Eastern Europeans having arrived prior to 2004 [Becker et al., 2016]. The number of these Polish-born migrants increased by a factor of seven, and the number of Eastern Europeans in the UK by a factor of five, such that the number of AC-12 migrants living in England represented approximately 2% of the English population in 2011, reaching 1,085,351 inhabitants.

5.3 Sampling frame and data sources

In this section, we first describe both the data on local authority expenditure and revenue and the migration data we use for our subsequent analyses, in subsections 5.3.1 and 5.3.2 respectively. We then describe additional data sources we draw on to obtain control variables and for the analyses of mechanisms that may explain the obtained results in section 5.3.3.

5.3.1 Local authority expenditure and revenue

Detailed panel data on local authority public finances is available from the Chartered Institute of Public Finance and Accountancy (CIPFA) website which gathers local authority budgeted expenditure and revenue data. We use this data set to identify exactly how the 116 single-tier authorities and 27 authorities operating under a two-tier regime allocated their funds between all different spending areas annually. We obtain these data for the period from 2000 to 2015 such that our sample consists of a total of 143 local authorities observed over a 16-year period. Table 5.2 summarises our main expenditure related dependent variables and also expresses these as expenditure shares for comparability.

Variable	Mean	Std. Dev.	N
Total expenditure per capita (pc)	1450.07	330.32	2028
Central government transfer (pc)	1086.04	335.83	2028
Council tax required (pc)	364.03	80.90	2028
Social care expenditure (pc)	375.15	102.68	2028
Education expenditure (pc)	734.42	189.07	2028
Other expenditure (pc)	340.66	94.51	2028
Share spent on social care	0.258	0.039	2028
Share spent on education	0.506	0.057	2028
Share spent on other items	0.236	0.037	2028

Table 5.2: Summary statistics outcomes

The main variables we construct from this data set are our measures of expenditure and revenue per capita: total service expenditure, central government transfers and the required locally generated revenue to balance the budget. We also use this data set to construct our set of expenditure measures on education, social care and aggregate measures of spending on other items.

5.3.2 Migration data

We draw our yearly data on population for each local authority from a special license of the UK Annual Population Survey (APS) between 2000 and 2015. This sample is obtained by aggregating waves of the Labour Force Survey (LFS) and the Labour Force Survey Boosts for England.[2] We refer to this data set as LFS/APS. We are then able to compute immigration figures of AC-12 migrants as well as natives, EU-15 and the rest of the world based on their country of birth. There are approximately 350,000 individuals per wave, making the LFS/APS the largest annual household survey in the U.K.. Although the LFS/APS is more robust than estimates based on one single LFS wave, concerns regarding the accuracy with which this survey precisely measures the shares of immigrant population at the local authority level remain, especially in years preceding the 2004 shock when the AC-12 stock in England was relatively small. In our empirical strategy, we explain that the way we construct our main variable of interest does not require us to make use of the AC-12 migrant population pre-2004 in our regression analyses.

We further verify the robustness of our results by exploiting 2001 and 2011 British Censuses, which capture information for the entire British population by country of birth by local authority at these two points in time. Despite the 2011 Census' advantage of

[2]See [Cangiano, 2010] for more details on the APS.

reporting data on the self-reported year of arrival of different migrant groups for the years 2009, 2006, 2003, 2000 and 1990, concerns when using this data set remain. Differential rates of out-migration across local authorities by AC-12 migrants by year of arrival could mean that an imputed measure of migration stocks would have a poor relationship with the actual stock for years further in the past. We find a local authority level correlation between the number of foreigners from AC-12 migrants having arrived before 2000 as per the 2011 Census and the number of AC-12 migrants as reported by the 2001 Census of 0.77. Overall, this suggests the LFS/APS is the preferable data set. We nevertheless use the 2001 and 2011 Censuses to check our results for robustness.

In sum, given several missing values in CIPFA and LFS/APS data, our total sample consists of 2,028 observations spanning from 2000 to 2015 for 143 English local authorities.

5.3.3 Additional data sources

Our main specification measuring the effect of AC-12 migration controls for a number of local authority level time varying characteristics. First, we control for local area population obtained from CIPFA to account for potential scale effects in service delivery, especially regarding education. Since the effect of AC-12 on local redistribution may also be conditional of the existing composition and heterogeneity of the population, we further control for the share of EU-15 and non-EU migrants obtained from the LFS/APS [Alesina et al., 2019, Tabellini, 2020b]. Finally, we also control for local unemployment rates to account for the fact that local economic conditions are an important pull factor for labour migrants such as those originating from AC-12 countries.

We then construct additional dependent variables to understand the mechanisms driving our main results. These variables include information on the number of pupils per local authority, and the age structure of the population (both obtained from CIPFA). To test for a change in political preferences, we use yearly data on the political composition of local Councils compiled by "The Elections Centre" at Nuffield college, Oxford.[3] In England, local Councillors are elected in a staggered way for 4-year terms by the local population, with some Councils electing all of their Councillors at the same time and other Councils electing half or a third of their Councillors at each election.[4] In our analyses, we are particularly interested in the shares held by the Conservative and Labour

[3]The data can be accessed on http://www.electionscentre.co.uk.

[4]See https://www.gov.uk/understand-how-your-council-works/local-councillors-and-elections for details.

party, England's two main parties of government. The Conservative party is traditionally regarded as representing less redistributive "small government" platforms during our period of study as highlighted by [Fetzer, 2019] who studies the impact of austerity measures conducted by the Conservative government in the 2010s on Eurosceptic attitudes and [Taylor-Gooby, 2013] who shows that an "analysis of manifestos for the two main UK parties from 1987 to 2012" shows two patterns: First, Labour manifestos score higher than Conservative manifestos on references to social justice and pro-welfare content. Second, Conservative interest in social justice as equality or redistribution is limited, and is "virtually non-existent for the 1987, 2001 and 2005 elections" (p.36).

We further substantiate the strong differences between the Conservative and Labour party on redistribution in figure 5.1.

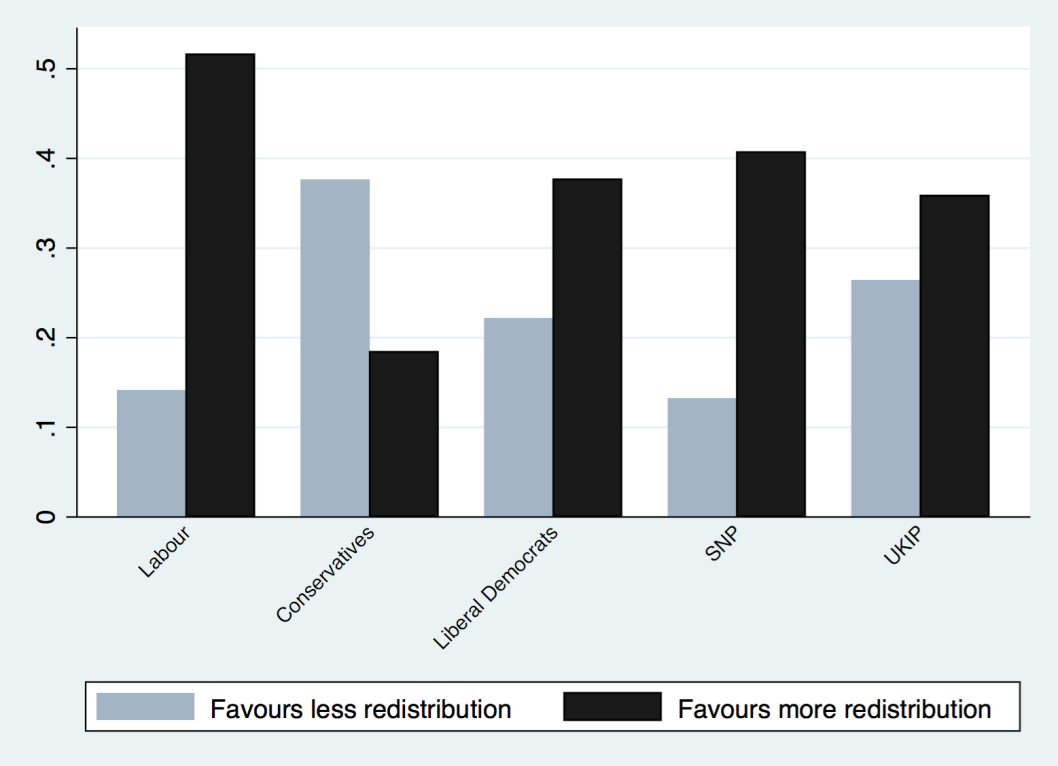


Figure 5.1: Redistribution preferences among voters of major British parties

Note: Figure based on data from the British Election Study (Wave 4, March 2015). The redistribution preference are calculated from the question "Should the government redistribute income?" to which respondents could respond on a scale from 0 ("make much greater effort") to 10 ("be much less concerned"). Respondents answering 0,1,2 or 3 are categorised as "favours more redistribution". Respondents answering 7,8,9 or 10 are categorised as "favours less redistribution". N=2840.

The figure based on data from the March wave of the British Election Study 2015 shows that more than 50% of Labour voters favour more redistribution of income, whereas

that share stood at less than 20% among Conservative voters, who tend to favour less redistribution. Voters of Liberal Democrats, SNP and UKIP fall in between.

Unfortunately, the Council composition dataset aggregates information on the Council seat share of the UK Independence Party (UKIP) into an "other" category. Founded in 1991 and until the 2016 Brexit referendum, UKIP was essentially a single-issue party campaigning for an exit from the EU and made strong gains in European elections at the expense of the Conservative Party over our sample period [Ziebarth, 2020, Becker et al., 2016, Fetzer, 2019]. However, it is is not clear that more votes for UKIP are a signal of preferences towards less redistribution, as highlighted by figure 5.1 . Nevertheless, greater vote shares for UKIP in the face of AC-12 migration still suggest greater distaste for migration, which could potentially result in changes in local Councillors' spending decisions. We therefore additionally match our database with information on all local election results from [Fetzer, 2019], which contains election results of almost all parties including UKIP. Finally, we also use information from the second quarter of each LFS wave as well as NHS data to build local authority level measures of internal inflows and outflows.

We note that not all data we use for our analyses of mechanisms is available for the entire 2000 to 2015 observation period. We summarise data availability, sources and the measures we construct in table 5.11.

5.4 Empirical strategy

In this section, we first lay out our main empirical strategy in subsection 5.4.1. We then turn to the main empirical issue in our setting at hand, the potentially endogenous location choice of migrants within England in subsection 5.4.2.

5.4.1 Main specification

The AC-12 migration wave to England, and the U.K. more generally, was a grand-scale natural experiment due to its large and unexpected nature. To capture its magnitude by local authority, we closely follow [Becker et al., 2018] and build the following shock

measure:[5]

$$Shock_{c,t} = \frac{AC12migrants_{c,t}}{LApop_{c,t}} - \frac{AC12migrants_{c,2001}}{LApop_{c,2001}} \tag{5.1}$$

where $AC12migrants_{c,t}$ represents the number of AC-12 migrants in local authority c in year t and $LApop_{c,t}$ local authority c's population in year t as per the APS while $AC12migrants_{c,2001}$ and $LApop_{c,2001}$ represent the AC-12 stock and total population in 2001 as per the 2001 Census. The shock measure thus represents the change in percentage points of the share of AC-12 migrants living in a given local authority relative to the share of AC-12 migrants in 2001. As argued earlier, the APS might not capture the share of AC-12 migrants by local authority pre-2004 well because of their low numbers. We therefore use the more precise 2001 census as our benchmark share of AC-12 migrants. Furthermore, we never use APS information regarding the AC-12 population pre-2004 as highlighted in our main empirical specification:

$$Y_{c,t} = \alpha_c + \beta_t + \gamma Post_{2004} * Shock_{c,t} + \delta X_{c,t} + \epsilon_{c,t} \tag{5.2}$$

In equation 5.2, $Post_{2004}$ is a dummy equal to 1 after 2004 such that its interaction with $Shock_{c,t}$ identifies a treatment on the treated effect of AC-12 migration on the outcome measures. $X_{c,t}$ represents our vector of controls which includes the local share of EU-15 migrants, the share of non-EU migrants, the local authority population and the local unemployment rate. We further include year fixed effects β_t and local authority dummies α_c to account for unobservable year-specific variation common to all local authorities and time-constant local authority level variation respectively. Due to the spatial nature of our data, we cluster the error term $\epsilon_{c,t}$ at the local authority district level in all analyses to correct standard errors for within local authority correlations [Bertrand et al., 2004]. Our estimation is therefore akin to employing a two-way fixed effects panel model that combines features of a continuous Difference-in-Differences and event study as we set the pre-accession AC-12 stocks to zero. We do this due to the to the strong differences between AC-12 pre-and post 2004 and expect a gradual divergence of treated and untreated local authorities in the post treatment period, as not all migrants moved to the UK at the start of our treatment period.

Table 5.3 describes our main variable of interest - AC-12 migration shock - and controls' summary statistics.

[5]The main contrast in the construction of our shock measure is that we use the LFS/APS instead of only using the 2001 and 2011 Censuses to measure changes in migrant shares.

Variable	Mean	Std. Dev.	N
AC-12 migration shock	0.011	0.015	2028
Share EU-15 migrants	0.027	0.023	2028
Share non-EU migrants	0.101	0.097	2028
Unemployment	0.068	0.026	2028
Population	358010	270538	2028

Table 5.3: Summary statistics AC-12 shock and controls

Local authorities' average population stands approximately at 358,000 inhabitants with the average increase in the share of AC-12 migrants representing a 1.1 percentage point increase in the share of a local authority's population compared to the 2001 baseline.

5.4.2 Endogenous location choices

The main concern for causal interpretation of γ is the endogenous location choices of AC-12 migrants. Despite the exogenous nature of the shock at hand, immigrants were not randomly allocated across local authorities and sorting into local authorities might be endogenous to migrants' underlying preferences for redistribution or other unobserved characteristics that determine both migration and trends in local authority spending. We first address this concern by presenting evidence in favour of the common trends assumption by showing that all measures of interest did not move systematically differently between affected and unaffected local authorities prior to 2004 in 5.5 and 5.6.

We then show that our results are practically unchanged when matching local authorities affected by AC-12 migration inflows to unaffected local authorities using propensity score matching. To do so, we follow an approach similar to Becker et al. [2018] and first apply a corrected Akaike Information Criterion (AICc) algorithm to a large set of local authority characteristics we gather from the 2001 Census to find the best predictors for AC-12 migration inflows and then match our sample based on these variables (with replacement). While this approach reduces our sample size, it mitigates remaining concerns that destination choices were not picked at random and should be treated as complementary to our main difference-in-differences results.

Finally, we proceed to showing the results when using a shift-share "Bartik" instrumental variable approach based on historical settlement [Bartik, 1991]. The problem of endogenous sorting of migrants has indeed often been tackled by using such an approach pioneered by Card [2001]. It is based on the premise that immigrant networks are an important determinant of locational choices and allows to identify local average treatment

effects of current migration inflows induced by these historical settlement patterns. In the context of the UK, this instrument has been used by Bell et al. [2013], Sá [2015] and Giuntella et al. [2018] in their studies on the impact of migration to the UK and its effect on crime, house prices and hospital waiting times respectively. The validity of such an instrument relies on the assumption that the past settlement of immigrants is uncorrelated with changes in economic outcomes between local authorities prior to 2003. Thus, it assumes that immigrant settlement patterns years before 2004 are only correlated with measures of local authority spending patterns though their effect on post-2004 inflows.

While we show that our results are stable when using such an approach based on 1991 settlements of AC-12 migrants, several problems have been identified in shift-share instrumental variables. Jaeger et al. [2018] first show that such instruments run the risk of conflating the short- and long-run responses to immigration shocks if the spatial distribution of immigrant inflows is stable over time. Second, Lee et al. [2020] find that current practice using instrumental variables typically relies on the first-stage F-statistic exceeding a threshold of 10 as a criterion for showing instrument relevance and trusting t-ratio inference yield an anti-conservative test. In addition to these concerns, we believe this approach is potentially less relevant when applied to AC-12 migrants. Becker et al. [2018] point out that the historical distribution of Central and Eastern European may not capture subsequent AC-12 migration inflows well because of the stark differences between the two waves. The authors argue that migrants from Poland – the country where the largest share of AC-12 migrants originate from - who resided in the UK in 2001 mainly consisted of people of pension age, having lived in the UK since the second world war as remnants of the Polish Free Army, or of migrants who had entered before 1991 for graduate studies under high-skill visas. Thus, the baseline distribution of Central and Eastern European migration into the UK in 2001 would act as a weak proxy for subsequent migration flows or could pick up very specific parts of these migration flows. This is confirmed when conducting standard first-stage relevance tests but also by comparing the distribution of AC-12 migrants in 1991 and in 2015 as highlighted in figure 5.2. The figure shows that recent AC-12 migration was more heterogeneously distributed across England than in 1991, when a concentration in London and its surroundings was more clearly visible.

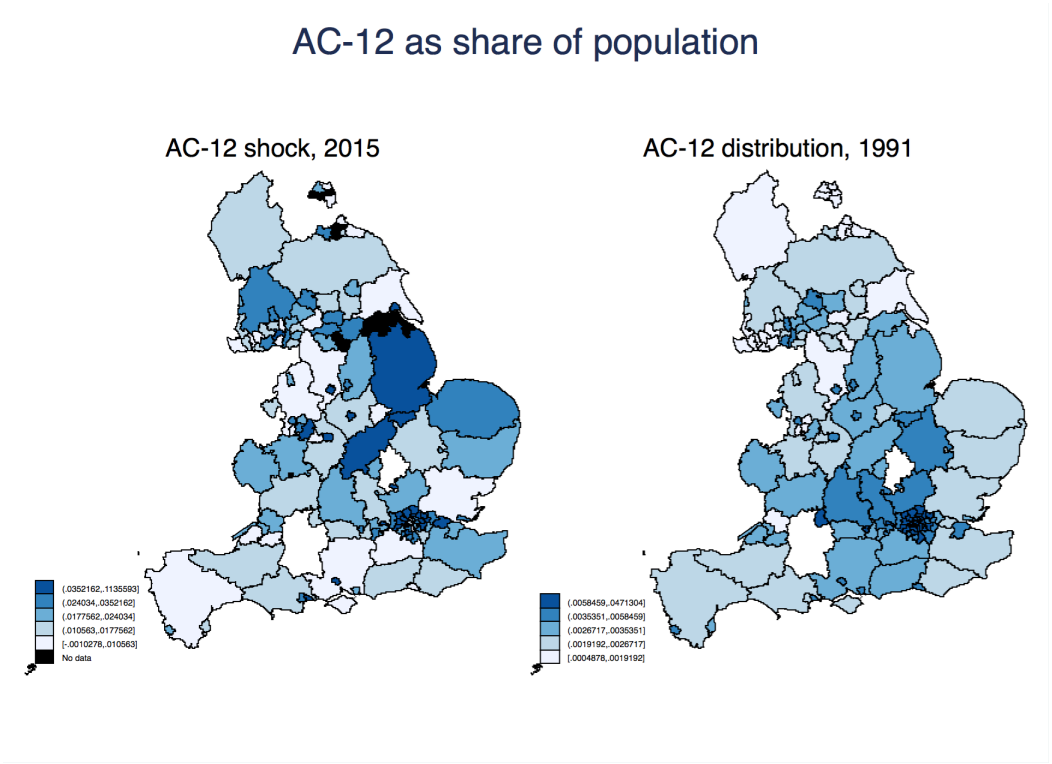


Figure 5.2: Contemporaneous AC-12 shock and historical distribution

5.5 Results

In this section, we turn to the results and split our analyses into two parts. Subsection 5.5.1 provides evidence in support of the identifying common trend assumption. Subsection 5.5.2 then shows the main results.

5.5.1 Pre- and post trends

In this subsection, we analyse the main outcomes of interest before and after the opening of UK borders to AC-12 migrants in 2004. Figure 5.3 shows that the share of AC-12 migrants indeed gradually picked up after 2004, starting from a level close to zero before the accession of the AC-12 countries.

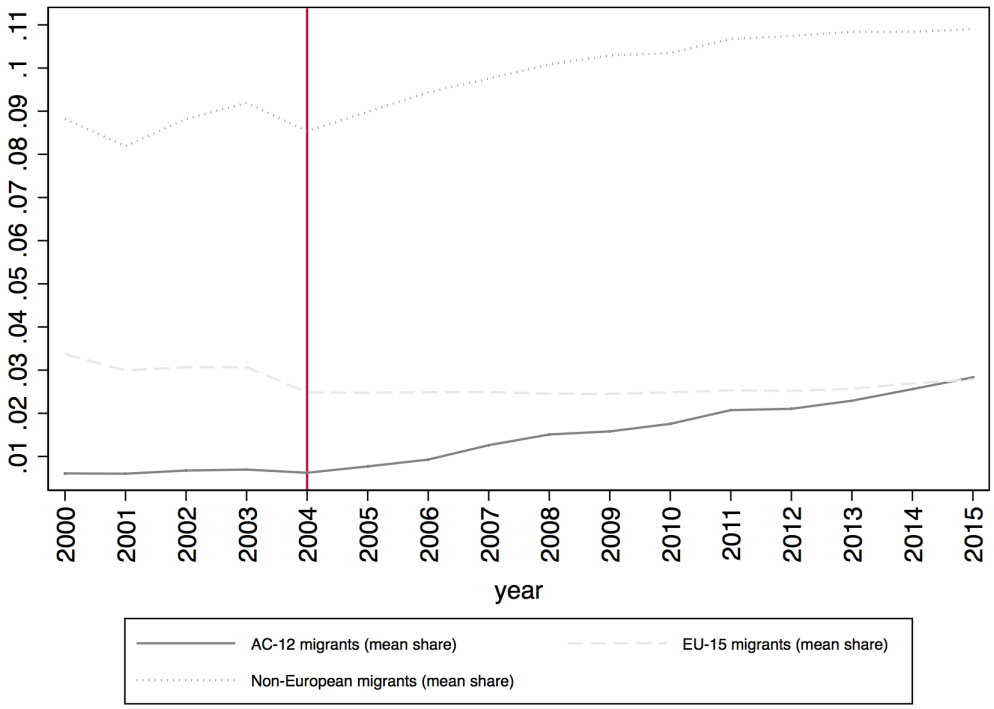


Figure 5.3: Migration to the UK by region of origin

The mean annual local population share of AC-12 born migrants in English local authorities fluctuated between 0.6% and 0.7% in the pre-accession years and then gradually increased to 2.8% in 2015. In relative terms, the average English local authority saw their AC-12 migrant population rising by more than 400% over an 11 year period.

Figure 5.3 further shows an increase in the local authority mean population share of non-EU migrants between 2004 and 2015 while the EU-15 population share remained stable on average. One of the concerns in our empirical analysis when focussing on one migrant group specifically is that location choices of AC-12 migrants may coincide with inflows of other migrant groups. In appendix 5.C we show that the spatial distribution of non-EU and EU-15 migrant inflows does not correlate visibly with the AC-12 shock visualised in figure 5.2. We nevertheless control for these migrant shares in our preferred specification to adequately capture the diversity of the local area population as discussed in section 5.4.1.

To test for pre- and post-trends in our main outcome variables of interest, we define local authority districts that were in the top 25% (Q4) of the 2015 migration shock distribution as treated ("affected") . We further define all local authority districts that are within the bottom 25% (Q1) of the 2015 migration shock distribution as untreated ("unaffected"). Formally, we define

$$\tau_c = \begin{cases} 1, if & AC12Shock_{2015,c} > Q3(AC12Shock_{2015}) \\ 0, if & AC12Shock_{2015,c} < Q1(AC12Shock_{2015}) \end{cases} \tag{5.3}$$

The categorisation is motivated by the distribution of the AC-12 shock measure in our final observation year in 2015 shown in figure 5.4.

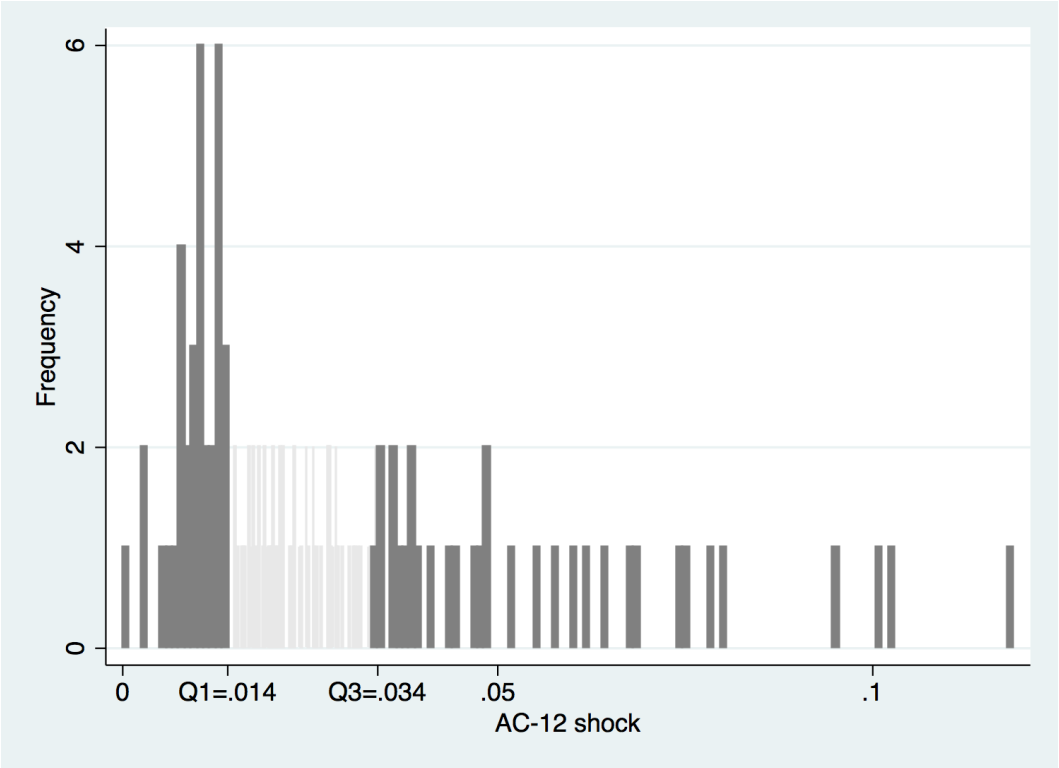


Figure 5.4: Distribution of AC-12 shock in 2015

Local authorities in the bottom quartile of the AC-12 shock distribution only saw their population share of AC-12 migrants rising by between 0.1% and 1.4%. The frequency distribution then has a long right tail with AC-12 migrant shares rising to between 3.4% and 11.9% in local authorities in the top quartile.

Figures 5.5 and 5.6 show the trends in the main outcome variables of interest, with the solid lines corresponding to regions most heavily affected by AC-12 migrant inflows and dashed lines indicating the least affected regions.

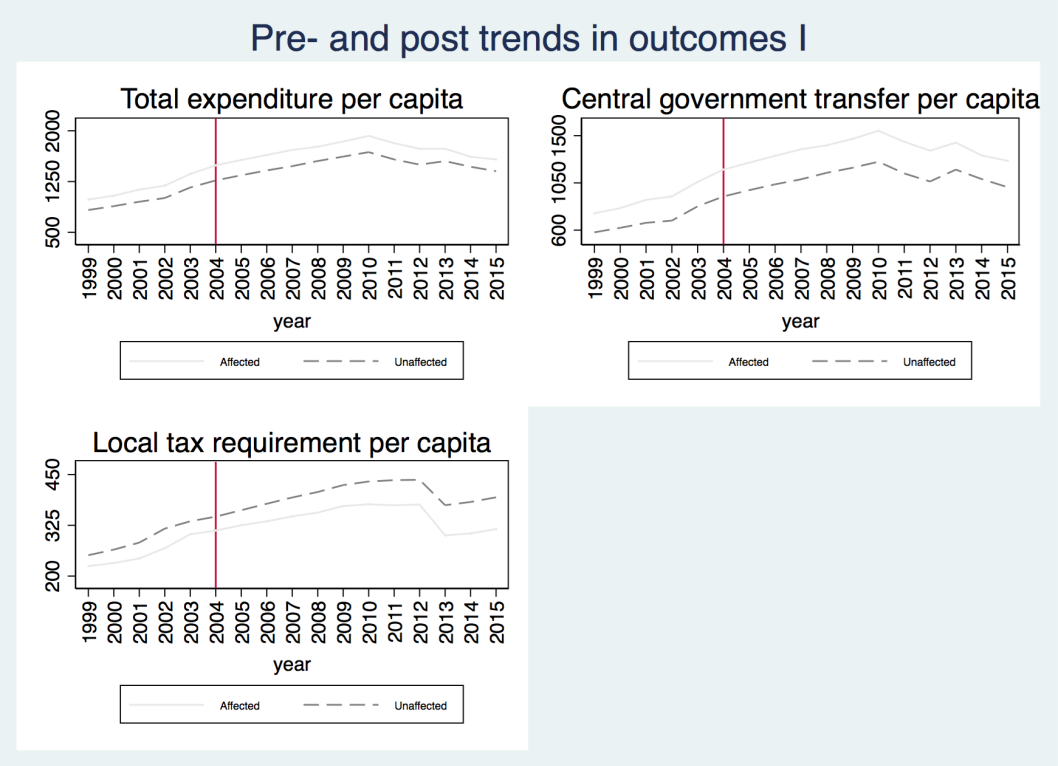


Figure 5.5: Pre- and post trend I

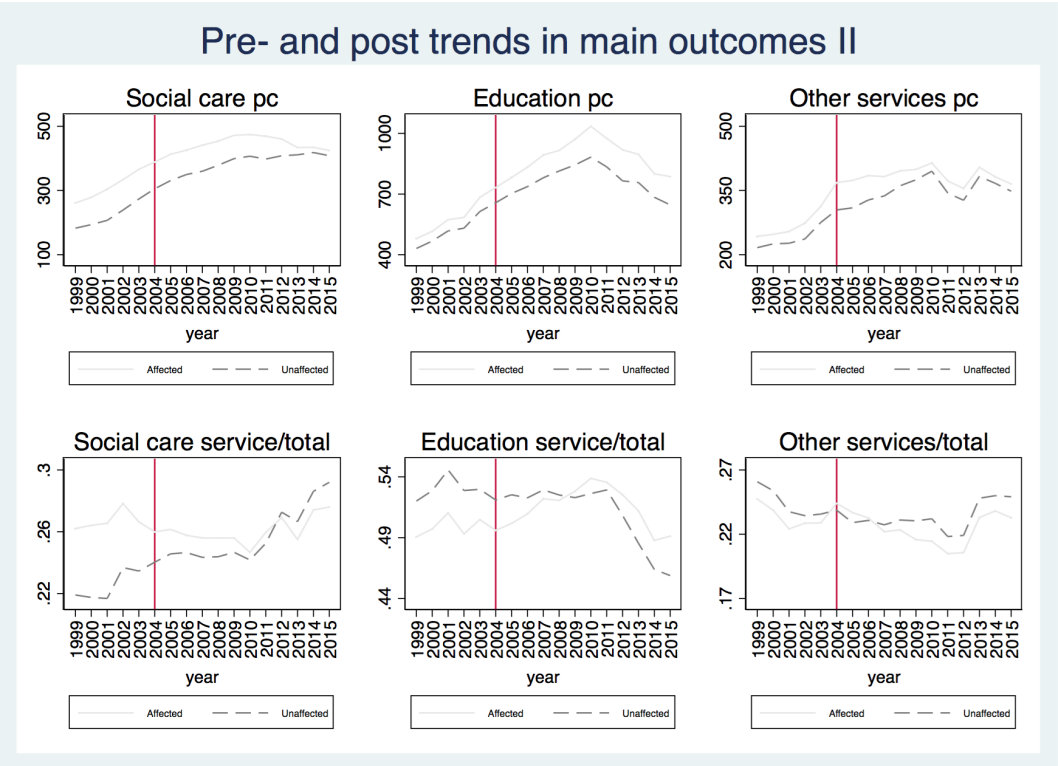


Figure 5.6: Pre- and post trend II

We first note that pre-trends from 1999 to 2003 are parallel in all outcomes, validating the difference-in-differences research design. We detect no visual changes in the total expenditure per capita gap post accession (figure 5.5, upper left panel). Affected and unaffected local authorities continue to move in parallel after the onset of treatment. The parallel trend continues beyond 2010, when strict austerity measures were imposed on local authorities, a topic we turn to in greater detail in section 5.6.1. The lower left and upper right panel of figure 5.5 suggest a marginal effect of AC-12 migrant inflows on the funding sources of local authorities. Taken together, the figures suggest that in affected areas, less funding was raised locally and slightly more funding came from central government sources following the inflow of AC-12 migrants, leading to the observed unchanged total per capita spending in an institutional setting where local authorities have to balance their annual budgets.

Figure 5.6 then turns to the spending mix. Social care expenditure per capita starts to rise significantly less in affected local authority areas when AC-12 migrants start to immigrate in 2004. The gap most visibly declines in the years from 2010 onwards (top right panel). The change in trends becomes even more visible when graphing social care as a share of overall local authority spending, where the social care expenditure share in unaffected local authorities passes the share spent in affected local authorities in 2012 (bottom right panel). Similarly, trends in education expenditure per capita start to diverge after 2004, with local authorities most heavily affected by AC-12 migrant inflows spending relatively more on education in per capita terms (top centre panel). The share spent on education in total spending surpasses that of unaffected areas in 2009 and the gap continues to widen up to the final year of the observation period (bottom centre panel). Other expenditure items expressed in per capita terms show slightly more noise but tend to converge after 2004, with affected areas experiencing a small relative decrease (figure 5.6, upper left and bottom left panel).

In sum, three main findings emerge from this first descriptive analysis. First, total expenditure per capita remained unchanged following AC-12 migrant inflows. Second, the funding mix slightly shifted away from locally raised budget towards central government transfers. And third, AC-12 migrants shifted the expenditure mix away from social care expenditure towards education expenditure.

5.5.2 Main results

In this subsection, we turn to the regression results based on the empirical specification introduced in section 5.4. Table 5.4 shows the results of the effect of AC-12 migration on local spending and revenue sources. For the sake of clarity, we only show the coefficient of interest in the following tables. A regression table also showing the coefficients estimated on our main control variables is shown in table 5.14 of appendix 5.A. We will refer to the specification with all controls as our preferred specification.

	Total expenditure pc		*Central government transfer pc*		*Locally raised budget pc*	
	(1)	(2)	(3)	(4)	(5)	(6)
AC-12 shock	460.44	525.65	1014.67**	976.97**	-554.22***	-451.33***
	(459.70)	(457.89)	(483.08)	(472.25)	(96.65)	(87.04)
N	2028	2028	2028	2028	2028	2028
R^2	0.879	0.890	0.821	0.834	0.819	0.826
Time FE	Year	Year	Year	Year	Year	Year
Model	DiD	DiD	DiD	DiD	DiD	DiD
Full set of controls	No	Yes	No	Yes	No	Yes
Sample	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS

Standard errors clustered at the local authority level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: Outcomes expressed in per capita terms (pc). The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population. LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey. The observation period is 2000 to 2015.

Table 5.4: Effect of AC-12 migration on local spending and revenue

Column (1) and (2) show that the association between the AC-12 shock measure and total service expenditure per capita is positive, albeit not significant at any conventional level. Similar to the descriptive analysis in subsection 5.5.1, local authority revenue sources were indeed impacted by AC-12 migration: Column (3) indicates that a 1 percentage point increase in the AC-12 shock over the 2001 baseline shares is associated with an annual increase of approximately GBP 10 in central government transfers per capita to affected local authorities ($p<0.05$). The result is remarkably stable to the inclusion of controls for local labour market conditions, population and other migrant groups (column 4). The results of column 5 and 6 show that a percentage point increase in the AC-12 migrant population share over the 2001 baseline is associated with a GBP 5 decline in the locally raised budget per capita ($p<0.01$). The coefficients are again stable when adding controls variables. It is worth noting that the coefficients estimated on EU-15 and non-EU migrant shares point into the same direction. These coefficients shown in columns (1), (2) and (3) of table 5.14, appendix 5.A, indicate that AC-12 migration affected central government transfers in a similar manner compared to migrants from other regions.

Table 5.5 then turns to the results of the effect of AC-12 migration on local expenditure by area of spending.

	Social care expenditure pc		*Education expenditure pc*		*Other expenditure pc*	
	(1)	(2)	(3)	(4)	(5)	(6)
AC-12 shock	-477.89**	-414.47**	1214.59**	1179.50**	-279.10	-238.21
	(219.71)	(204.35)	(574.23)	(564.80)	(178.60)	(152.78)
N	2028	2028	2028	2028	2028	2028
R^2	0.771	0.790	0.750	0.757	0.641	0.654
Time FE	Year	Year	Year	Year	Year	Year
Model	DiD	DiD	DiD	DiD	DiD	DiD
Full set of controls	No	Yes	No	Yes	No	Yes
Sample	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS

Standard errors clustered at the local authority level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: Outcomes expressed in per capita terms (pc). The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population. LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey. The observation period is 2000 to 2015.

Table 5.5: Local authority spending by area

The results indicate that the AC-12 migrant shock indeed changed the local authority spending mix. The results of our preferred specification of column (2) show that a 1 percentage point increase in the shock measure is associated with an approximate GBP 4 decrease in annual per capita spending on social care (p<0.05). Column (4) shows that education spending per capita increased by approximately double that magnitude (p<0.01). The increase in AC-12 migrant shares further had a small negative effect on other expenditure items, albeit estimated effects are not significantly different from zero at any conventional statistical level. In comparison, changes in migrant shares from other regions of origin had less impact on the expenditure mix. Column (4) of table 5.14, appendix 5.A, suggests that the positive association of EU-15 migrant shares and total expenditure per capita (column 1, 5.14) is the result of an increase in additional spending on education associated with this particular migrant group. Neither EU-15 migrants nor non-EU migrants appear to be associated with changes in social care expenditure (column 4, table 5.14).

In sum, three main results emerge. First, the inflow of AC-12 migration does not show a statistically significant association with total service expenditure per capita. Second, AC-12 migration did impact on local authority revenue sources: They are associated with an increase in funding received from the central government but decreased locally generated income. Finally, AC-12 migration decreased social care spending per capita and increased education expenditure per capita, while other expenditure items remained

largely unchanged. We note that these findings do not allow for making definite statements about the mechanisms at hand: A reduction in the mostly discretionary means-tested social care services and the relative reduction of locally generated revenue in response to AC-12 inflows could indicate both a shift in local preferences towards less redistribution but could also capture changes in local service demand brought about by distinct socio-economic characteristics of new-arrivals, with repercussions for necessary funding. We analyse these mechanisms in more detail in section 5.7, after testing the robustness of our results in section 5.6.

5.6 Robustness tests

In this section, we test the robustness of our main results along a range of dimensions. Subsection 5.6.1 excludes all years post 2010 to account for the potentially endogenously austerity measures imposed on local authorities in 2010 as well as the expiration of the UK's Worker Registration Scheme (WRS) in 2011. If, for example, austerity measures were imposed on the central government level such that more diverse areas were more heavily affected, these would not necessarily reflect a local change in preferences for redistribution. In subsection 5.6.2 and 5.6.3 we further address potential endogenous sorting of AC-12 migrants by showing the results of a matching and an instrumental variable approach. These robustness tests thus account for the risk that AC-12 migrants potentially chose their destinations within England based on observed or unobserved local authority characteristics that could be correlated with our outcome measures. Finally, in subsection 5.6.4 we use UK Census data instead of the more volatile Annual Population Survey data as a more accurate but static measure of AC-12 migration.

5.6.1 Excluding austerity years and the end of WRS

Austerity measures imposed on local authorities by the central government dramatically reduced transfers to local authorities from 2010 onwards, putting pressure on local authorities to raise budget locally. These measures have been shown to have uneven effects on local authorities across England, mostly affecting urban areas, the north of England, parts of the East and Cornwall [Gray and Barford, 2018]. If these geographically heterogeneous budget cuts imposed by the central government correlate with the different intensity of AC-12 migration inflows for either political or institutional reasons unrelated

to *local* preferences for redistribution, they could potentially taint the estimated effects.

A further important institutional change coinciding with the introduction of austerity measures is the expiration of the UK's Worker Registration Scheme (WRS) in 2011. The WRS stipulated that AC-8 migrants - that is, migrants who originate from the 2004 Central and Eastern European EU joiners - could claim out-of-work benefits and tax credits on the same grounds as other EEA nationals only after being registered to the WRS and in continuous employment for 12 months. [Giua, 2020] shows that the expiration of the WRS had a positive impact on the probability of claiming out-of-work benefits by these migrants. Access to some of the local authority services such as social care services only required registering on the scheme, mitigating the concern of a pick-up in demand after the expiration of WRS[Kofman et al., 2009]. However, the pick-up in claims of out-of-work benefits by AC-8 migrants post 2011 may have had a second-order effect on demand for means-tested local authority services due to its direct effect on income. A general link between local unemployment rates and demand for means-tested social care services is indeed visible in our regressions (column 4, table 5.14, appendix 5.A). To account for these two important regime changes, we therefore test our results for robustness when excluding post-2010 years.

Table 5.6 shows the results estimated in our preferred specification when excluding all years post 2010.

Table 5.6: Effect of AC-12 migration on local spending and revenue - pre-austerity

	(1) Total expenditure pc	(2) Central government transfers pc	(3) Locally raised budget pc	(4) Social care pc	(5) Education pc	(6) Other pc
AC-12 shock	787.90*	1027.55**	-239.65**	-184.75	1309.79***	-336.15
	(463.06)	(471.63)	(104.79)	(179.80)	(435.48)	(206.63)
N	1343	1343	1343	1343	1343	1343
R^2	0.953	0.925	0.872	0.874	0.933	0.734
Time FE	Year	Year	Year	Year	Year	Year
Model	DiD	DiD	DiD	DiD	DiD	DiD
Full set of controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Notes: Outcomes expressed in per capita terms (pc). The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population. LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey. The observation period is 2000 to 2010.

The results broadly confirm those of the longer panel shown in tables 5.4 and 5.5. All coefficients are of the same sign in the early years of AC-12 inflows, suggesting that neither the expiration of the WRS nor the centrally imposed austerity measures fundamentally changed the effect AC-12 migrants had on local authority spending and revenue. Two observations are nevertheless noteworthy. First, all results are slightly less precisely estimated, likely a consequence of the significantly smaller sample size. Second, the coefficient estimated on social care spending per capita is smaller than in the full sample (column 4). One possible reasons for the amplification of observed changes in social care spending during the austerity years is a progressive adaptation by local authorities to changing demographics. A second possible explanation is a delay in the change in local preferences for redistributive polices. On the other hand, the more gradual increase in the demand for education services in local authorities more heavily affected by the migration shock is in line with the observation that many AC-12 migrants were of child-bearing age when they arrived to the UK, or required additional spending due on language schooling. We analyse these hypotheses in more detail in section 5.7.

5.6.2 Matching

The pre-trends we present in figures 5.5 and 5.6 show that AC-12 migrants did not systematically sort into local authorities based on trends in local spending or revenue regimes. However, a concern remains that destination choices were not picked at random: If AC-12 migrants moved into local authorities of specific underlying unobserved characteristics, these characteristics could have medium- to long-term consequences for local spending and revenue patterns that are then picked up by our estimates. For example, since AC-12 migrants were primarily labour migrants of a distinct skill profile, their destination choices were likely based on labour demand from specific industries; if the presence of these industries was linked to future local economic development in local authority areas, the association of AC-12 migrant shares and local area spending may be spurious.

In this subsection, we address this concern by matching affected and unaffected local authorities based on a wide range of local authority characteristics we draw from the 2001 UK Census. The local area characteristics we use for the matching procedure range from local authority expenditure and revenue measures to local household wealth indicators, local industry composition and demographics. A full list of variables is provided in appendix 5.D. The simple matching-with-replacement estimation proceeds in three steps. Similar to our pre and post-trend analysis, we first divide local authorities into affected

(=treated) and unaffected (=untreated) based on the increase in AC-12 migrants they experienced between 2004 and 2015. To increase the pool of potential matches, we define the top 50% receiving areas as affected and the bottom 50% of receiving areas as unaffected for the sake of the matching exercise. In a second step, we then use a simple stepwise Akaike's corrected information criterion (aicc) to select variables that best predict whether or not a local authority was treated, balancing an increase in the goodness-of-fit against the additional information required to achieve it [Cavanaugh and Neath, 2019]. The variables selected by the algorithm are highlighted in the full list shown in appendix 5.D. In summary, AC-12 migrants were most likely to migrate into local authority areas with relatively larger manufacturing, hotel, health, fishing, financial and domestic work industry shares. Further variables that predict AC-12's propensity to migrate into specific areas include household shares deprived in one dimension and the share of social housing provided by the local Council. In the final step, we then use propensity score matching to match treated and untreated local authorities based on the selected variables.

Table 5.7 shows the results of the matching regressions.

Table 5.7: Effect of AC-12 migration on local spending and revenue - matching

	(1) Total expenditure pc	(2) Central government transfers pc	(3) Locally raised budget pc	(4) Social care pc	(5) Education pc	(6) Other pc
AC-12 shock	664.77	1187.73**	-522.96***	-402.99*	1475.97**	-404.50**
	(520.56)	(539.25)	(115.50)	(233.09)	(616.92)	(174.83)
N	1295	1295	1295	1295	1295	1295
R^2	0.883	0.826	0.808	0.771	0.752	0.651
Time FE	Year	Year	Year	Year	Year	Year
Model	DiD (matched)	DiD (matched)	DiD (matched)	DiD (matched)	DiD (matched)	DiD (matched)
Full set of controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Notes: Outcomes expressed in per capita terms (pc). The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population. Local authorities in the top 50% of AC-12 migrant receiving areas in 2015, measured by the AC-12 shock, are defined as affected and matched with the bottom 50% of AC-12 migrant receiving areas based on 2001 UK Census characteristics described in appendix 5.D. Matching conducted with replacement.LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey. The observation period is 2000 to 2015.

Matching local authorities based on their 2001 characteristics does not alter the main results. All estimated coefficients are of similar magnitude compared to the results based on the unmatched sample in tables 5.4 and 5.5. The results continue to indicate a shift in revenue sources away from locally raised budget towards central government transfers (columns 2 and 3) and a shift in expenditure from social care towards education (columns 4 and 5). The drop in local service expenditure outside education and social care in response is more precisely estimated in the matched sample ($p<0.05$). Overall, the results suggest that spatial sorting among AC-12 migrants based on observable local authority characteristics is not a large concern in the post-accession English setting. To further robustify this finding, we next turn to an an instrumental variable approach in subsection 5.6.3.

5.6.3 Instrumental variable approach

As discussed in subsection 5.4.2, a common way of dealing with endogenous destination choices is to instrument contemporaneous migration inflows by historical settlement patterns. This instrumental variable was pioneered by Card [2001] and is based on the premise that immigrant networks are an important determinant of locational choices [Altonji and Card, 2018]. In the context of the UK, such a "shift-share" instrument has been used by Bell et al. [2013], Sá [2015] and Giuntella et al. [2018] in their studies on the impact of migration to the UK and its effect on crime, house prices and NHS waiting times respectively. In our setting, the identifying assumption of such an instrument is that the historical settlement of AC-12 immigrants is correlated with post-2004 changes in local authority spending and revenue only through their effect on post-2004 inflows of AC-12 migrants. In this subsection, we use 1991 UK Census data to construct the instrument similar to Giuntella et al. [2018]. Specifically, we define $\lambda_{c,1991}$ as the share of AC-12-borns living in local authority c in 1991 and calculate:

$$\frac{\lambda_{c,1991} \sum AC12_t}{Pop_{c,t}}. \tag{5.4}$$

Thus, for each local authority c and year t, we multiply the 1991 share of AC-12 migrants residing in that local authority by the aggregate national level stock of AC-12 migrants in year t to project the new stock of AC-12 migrants. We then divide this number by the local area population $Pop_{c,t}$ to derive the projected shares.

We do not use the IV approach as our preferred specification for a number of reasons.

First, in the context of AC-12 migration, Becker et al. [2016] point out that it is unclear whether such a shift-share instrument captures the skill-composition of AC-12 migrants well. The authors argue that migrants from Poland – the country where the largest share of AC-12 migrants originate from - who resided in the UK in 1991 mainly consisted of people of pension age, having lived in the UK since the second world war as remnants of the Polish Free Army, or of migrants who had entered before 1991 for graduate studies under high-skill visas. This would mean that the baseline distribution of Eastern European migration into the UK in 1991 would act as a weak proxy for subsequent migration flows, a point confirmed by the relatively weak first-stage F-test of 5.93 shown in table 5.8. In the context of our research question, the distinct nature of historical inflows from AC-12 settlement in the UK raises an equally important question pertaining to the specifics of the local average treatment effects (LATE) such an instrumental variable identifies. The subset of contemporaneous AC-12 migrants pulled in by historical settlers from AC-12 countries is likely to be different from those AC-12 migrants who came to England for work. Thus, any LATE identified is potentially very different from the average treatment effect of AC-12 migrants on local authority spending researchers and policymakers are ultimately interested in.

With these caveats in mind, table 5.8 displays the results of the IV-regressions.

Table 5.8: Effect of AC-12 migration on local spending and revenue - IV approach

	(1) Total expenditure pc	(2) Central government transfers pc	(3) Locally raised budget pc	(4) Social care pc	(5) Education pc	(6) Other pc
AC-12 shock	3092.22	5002.80**	-1910.58**	-3293.37***	5853.21***	551.08
	(2052.34)	(2471.26)	(855.56)	(1215.52)	(2055.41)	(826.78)
N	2028	2028	2028	2028	2028	2028
Time FE	Year	Year	Year	Year	Year	Year
Model	DiD - IV	DiD - IV	DiD - IV	DiD - IV	DiD - IV	DiD - IV
Full set of controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS
First-stage F-test	5.93	5.93	5.93	5.93	5.93	5.93

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Notes: Outcomes expressed in per capita terms (pc). The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. DiD - IV refers to a difference-in-differences approach where the AC-12 shock variable is instrumented by the historical distribution of AC-12 migrants across England as defined in 5.4. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population.LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey. The observation period is 2000 to 2015. The F-test of the first-stage regression is the Kleibergen-Paap rk Wald F statistic. It is reported only for the post-shock period.

The LATE estimates obtained from the instrumental variable regressions confirm the main results. When interpreted causally, the results show that the subset of AC-12 migrants pulled in by historical networks had sizeable effects on both local authority revenue sources and the spending mix. All estimated coefficients are between five and eight times larger than those obtained in our preferred specifications of tables 5.4 and 5.5. We note that the difference in the magnitude of the obtained coefficients allows for two possible interpretations: First, the treatment-on-the-treated effects obtained from the difference-in-differences specification underestimate the effects AC-12 migration had on local authority budgets and service provision. A second interpretation is that the estimated LATE is identified based on a subset of migrants that differs significantly from the average group characteristics of AC-12. For the reasons outlined above, we believe that the second is more likely to hold true. We conclude that the estimates obtained in tables 5.4 and 5.5 are likely to be closer to the true effect AC-12 had on local authority service provision.[6]

5.6.4 Census data

Despite being relied on heavily in the UK-migration literature as a source of tracking local authority level migration stocks over time in the UK [Bell et al., 2013, Sá, 2015, Giuntella et al., 2018], the LFS/APS data we use as our primary data source has some disadvantages regarding its ability to accurately capture migrants of specific countries of origin on a granular geographical level [Cangiano, 2010]. The absolute number of AC-12 migrants or other migrant groups sampled in each local authority-year is sometimes too low to rely on in our empirical analyses, leading us to discard observation points when these counts fall below 10. In this section, we therefore confirm the main results using 2001 and 2011 Census data instead of the more volatile LFS/APS. Both censuses provide data at the local authority level of residents by country or region of birth while the 2011 Census provides information on the time of arrival of 2011 residents in two- or three-year brackets. As we do not have information within these year brackets, we calculate the stock of migrants as the last year of each bracket: For example, a migrant reporting to have entered a local authority district five to seven years ago in 2011 enters our stock calculation for the year 2006. This way, we obtain stock data for AC-12 migrants in every local authority district for the years 2000, 2003, 2006, 2009 and 2011. A disadvantage

[6] In theory, the relatively weak first stage F-test (F=5.93) may both lead to a bias of the estimated coefficient towards OLS -in which case the true coefficients would be even larger than those obtained - and an underestimation of the size of the standard errors around these point estimates Murray [2006].

compared to our main AC-12 shock measure is that all results are estimated exploiting year-on-year variation within local authorities only when stock data is available. The way the stock of residents is reported in the Censuses further means that it does not count migrants who arrived in England before 2011 and then left England before 2011 or who died before 2011. The stock of migrants reported for every year is therefore a lower bound estimate of the true migration inflow. The alternative shock measure is summarised in table 5.9.

Variable	**Mean**	**Std. Dev.**
AC-12 migration shock (alternative)	0.013	0.017
N	2028	

Table 5.9: Summary statistics

Table 5.10 then turns to the results using the alternative AC-12 migration shock measure based on Census data.

Table 5.10: Effect of AC-12 migration on local spending and revenue - Census data

	(1) Total expenditure pc	(2) Central government transfers pc	(3) Locally raised budget pc	(4) Social care pc	(5) Education pc	(6) Other pc
AC-12 shock (alternative)	1322.18***	1779.49***	-457.31***	-622.42***	2153.44***	-208.84
	(330.69)	(344.50)	(94.18)	(195.69)	(438.68)	(160.81)
N	2028	2028	2028	2028	2028	2028
R^2	0.891	0.837	0.825	0.793	0.767	0.654
Time FE	Year	Year	Year	Year	Year	Year
Model	DiD	DiD	DiD	DiD	DiD	DiD
Full set of controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Census	Census	Census	Census	Census	Census

Standard errors clustered at the local authority level in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Notes: Outcomes expressed in per capita terms (pc). The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population. Census refers to the 2001 and 2011 UK Census. The observation period is 2000 to 2015.

The results confirm the main results shown in tables 5.4 and 5.5. The estimated coefficients on total service expenditure per capita (column 1), central government transfers per capita (2), social care spending per capita (4) and education per capita (5) are by a magnitude of 1.5 to three times larger than those estimated in the baseline specification. All coefficients are slightly more precisely estimated. Due to the shortcomings of the Census data discussed above, we do not overinterpret these differences in effect size; however, the results show that imprecisely measured AC-12 migration stocks are unlikely to be the drivers of the association between AC-12 migration and local authority spending and revenue patterns.

5.7 Mechanisms

The results found in this paper document how a large migration shock translates into the provision of different public goods. In aggregate terms, our results suggest that the AC-12 migration shock only marginally affected the overall supply of public services in both pre and post-austerity years. On the revenue side, our results show a relative increase in central government transfers and a relative decline in locally generated revenue in local authorities affected by AC-12 migration inflows. We also observe important shifts in the types of public goods local authorities provide following the migration shock brought about by AC-12 migrants. Local authorities affected relatively more by migration inflows spent relatively less on means-tested social care services and significantly more on education services.

In this section, we investigate two important channels that might explain our observed results. On the one hand, the change in local authorities' populations stemming from AC-12 migration could lead to changes in redistribution via a change in preferences towards less redistribution or native flight lowering the tax base and local authority revenues. On the other hand, these changes could also reflect a more mechanical response. Migrants' have different socio-economic characteristics and different propensities to uptake services, which could lead to changes in redistribution without reflecting any migration-specific altruistic component. In this case, changes in redistribution from the rich to the poor do not discriminate between native-born versus foreign-born poor and simply reflect changes in the socio-economics characteristics of the population. Conditional on an institutional response by local authority councils, changes in the allocation of spending would then follow a mechanical response. We discuss both these channels separately in the following

subsections. While we acknowledge that both channels are non-exhaustive and could be at play simultaneously, we first discuss each of these channels and the hypotheses one can derive from them before arguing, through a number of tests, that the demographic channel better explains the observed changes in redistribution following AC-12 migration. For an overview of all variables we analyse as mechanisms and their source, data availability and construction, see table 5.11.

5.7.1 Migration and the altruism towards co-nationals: the preferences channel

The shift in expenditure from social care services towards education and the decline in the locally raised budget we identify could be interpreted as local authorities responding to a change in local preferences towards less redistribution following the migration shock. The link between migrant inflows and local preferences for redistribution could be a reflection of people from different groups disagreeing on the optimal amount and composition of local spending, or the dominant native group being less willing to redistribute towards non-co-nationals. While one has to be careful with using vote shares as an indication for an underlying preference for redistributing wealth or income in the population, we showed in figure 5.1 that in England, it is clearly the Conservative party that has support from voters less in favour of redistribution. We can thus test whether such a "demand" for less redistribution effect could explain our results by measuring the impact of AC-12 migration on the composition of local Councils and test whether the Conservative less pro-redistributive platform increased its seat share following a stronger migration wave. In this context, it is also worth noting that all EU (and Commonwealth) migrants to the UK have voting rights in local elections. Thus, an increase in the seats' share of Conservative Councillors, who represent the interest of voters with preferences for less redistribution, could also reflect a change in the aggregates populations' preferences if migrants exerted their voting rights.

A second more indirect mechanism through which distaste for outsiders could lead to less redistribution is a mechanism often referred to as "native flight" (see e.g. Cascio and Lewis [2012]) or "voting by feet". In addition to a direct reduction in demand for redistribution translating into changes in local spending, distaste for living near foreigners can also affect redistribution via natives moving into different local authorities. Such changes can indeed lead to a lowering of the tax base as the number of taxable properties decline, leading to a deterioration of local authorities' budgets and eventually a decline

in local authority government spending. To analyse such second-order internal migration in response to international inflows of migrants, we therefore report the impact of AC-12 migration on the number of taxable dwellings per capita (the tax base) as well as on the Band-D council tax rate, a lump-sum tax due to be paid annually by the dwelling owner. On average, this Band-D council tax stood at GBP 1278 over our observation period. To test the native flight hypothesis, we create two measures of internal migration. First, we draw on National Health Service (NHS) registration data. Getting basic access to free public health care in the UK requires registration with a local general practitioner. If individuals change their residence within the UK, they are required to provide their new general practitioner with any previous registration. The data thus allows to calculate both registrations and de-registrations for any given local authority and year. We then subtract internal migration inflows from internal migration outflows to construct a measure of net outflows of internal migrants at the local authority year level. Since the measure does not allow us to distinguish internal migrants by country of birth, we construct a second similar measure for UK native borns only, following [Giuntella et al., 2018]. Between 2000 and 2013, the second quarter survey of the UK-LFS contained a variable that asked respondents for their local authority of residence in the year prior to the survey, as well as their current local authority of residence. We use this information and calculate a "net internal native out-migration" variable for UK born individuals for all local authority years. Unlike the measure based on administrative NHS registration data, which is comprehensive, we normalise the UK-LFS data by the total number of respondents in each local authority-year to account for fluctuations in the sample size of the survey over time.

Table 5.11: Summary statistics - mechanisms

Variable	Mean	Std. Dev.	N	Source	Years available	Unit
Share of pupils in population	0.159	0.02	1755	CIPFA	2000-2013	$\frac{Pupils_{c,t}}{Population_{c,t}}$
Education per pupil	4664.09	1033.50	1755	CIPFA	2000-2013	$\frac{Education_{c,t}}{Pupils_{c,t}}$
Population aged >64	0.153	0.034	1881	CIPFA	2001-2014	$\frac{Population\,65+_{c,t}}{Population_{c,t}}$
Social care per population aged >64	2659.26	1365.57	1881	CIPFA	2000-2013	$\frac{Social\,care_{c,t}}{Population\,65+_{c,t}}$
Vote share labour party	0.398	0.267	2028	UK Elections Centre	2000-2015	$\frac{Council\,seats\,Labour_{c,t}}{Total\,seats_{c,t}}$
Vote share conservative party	0.371	0.243	2028	UK Elections Centre	2000-2015	$\frac{Council\,seats\,Labour_{c,t}}{Total\,seats_{c,t}}$
Net internal out-migration	198.57	2452.55	1840	NHS registrations	2002-2015	$Net\,outflow\,internal\,migrants_{c,t}$
Net internal native out-migration	0.007	0.049	1756	LFS	2000-2013	$\frac{Net\,outflow\,UK\,natives_{c,t}}{Population_{c,t}}$
Band D equivalent council tax	1278.14	225.49	2028	CIPFA	2000-2015	$Council\,tax_{c,t}$
Dwellings: Tax base per capita	0.34	0.06	2028	CIPFA	2000-2015	$\frac{Band\,D\,dwellings_{c,t}}{Population_{c,t}}$

5.7.2 Migration and changes to local authorities' socio-economic structure: the demographic channel

A second hypothesis that could explain the decline in local revenue and social care spending could simply be an institutional response to changes in impacted local authorities' underlying demographics. In our setting, local authority demographics are likely to have changed significantly in response to the distinct socio-economic characteristics of AC-12 migration inflows shown in table 5.1. These could have important implications for social care expenditure, which is predominantly consumed by elderly segments of the population. In 2004, the year of EU enlargement, the share of social care consumed by locals aged 65 and above stood at 40.1%. The observed decline in spending on social care could therefore simply reflect the per capita decline in demand for these services. We further recall that English local authorities only have upward discretion over education spending such that its relative increase could directly be linked to the rise in central government transfers that are channeled through local authorities [Phillips, 2018]. To investigate this institutional response, we begin by analysing changes in local demographics associated with AC-12 inflows by constructing a measure for the population share of pupils and a measure for the share of the population aged 65 and above to proxy demand for education services and social care services brought about by demographics respectively. In a second step, we then normalise changes in spending by these largest relevant consumer groups for both total spending on education and social care instead of using the previous per capita measure to test whether the observed changes were mechanical.

A second potential explanation for the uncovered association between AC-12 migration inflows and changes in local authority spending and revenue is a second-order effects brought about by an internal migration response to the new-arrivals: If, for example, the availability of relatively cheap labour creates local economic opportunities, this could draw in additional migrants from within the country. This, in essence, is the opposite of the "native flight" mechanism and can therefore be tested using the same internal migration measures as outcome variables as explained in subsection 5.7.3.[7]

[7] An additional mechanism we do not consider due to the lack of reliable data is migrants' propensity to demand local services, even once demographic differences are accounted for. For example, the "healthy migrant effect" could decrease the demand for social care services [Abraido-Lanza et al., 1999]. While data on social care referrals in England are available, these cannot systematically be divided into self-referrals and referrals by doctors. This is problematic since doctors are likely to refer their patients to social care services based on their understanding of availability of these services. A further issue with data on referrals is that they do not indicate the type (and thus, cost) of services requested.

5.7.3 The relative importance of the demographic channels: the easing off of pressure on social care stemming from AC-12 migration

Table 5.12 reports the results testing the hypotheses we have brought forward to discuss the relative relevance of the two main channels discussed above.

Table 5.12: Mechanisms

	(1) Share pupils	(2) Share 65+	(3) Education/pupils	(4) Social care/65+	(5) Labour seats	(6) Conservatives seats	(7) Net int. outflows	(8) Native outflows
AC-12 shock	0.15***	-0.26***	2947.91	5559.07***	0.65*	-0.54	-12811.18***	-0.31**
	(0.04)	(0.04)	(2684.36)	(1337.05)	(0.39)	(0.36)	(3968.12)	(0.13)
N	1755	1881	1755	1881	2028	2028	1840	1755
R^2	0.383	0.212	0.790	0.694	0.313	0.241	0.046	0.013
Time FE	Year	Year	Year	Year	Year	Year	Year	Year
Model	DiD	DiD	DiD	DiD	DiD	DiD	DiD	DiD
Full set of controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Notes: The outcome variables are defined in table 5.11. The number of observation varies due differences in data availability shown in detail in the same table. The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population. LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey.

We first turn to columns (1) to (4) to analyse the local demographic channel. AC-12 migration is strongly associated with an increase in the local population share of pupils ($p<0.01$; column 1) and is similarly associated with a decline in the local population aged 65 and above ($p<0.01$; column 2). The estimated coefficients indicate that a one percentage point increase in the share of local AC-12 migrants relative to the 2001 AC-12 population share changes the share of pupils in the population and the share of individuals aged above 65 by 0.15 percentage points and -0.26 percentage points respectively. Columns (3) and (4) then show the spending on education and social care relative to their main consumption groups. The coefficient estimated on education expenditure by pupil (3) shows that the large increase in education expenditure per capita we documented in the previous analyses is likely due to the change in underlying demographics associated with AC-12 inflows. The estimated coefficient is still positive, but no longer differs from zero at any conventional statistical level. In column 4, we show that the decline in social care service spending per capita associated with AC-12 inflows disappears when spending is calculated as a share of the main recipient group. In fact, increases in the local AC-12 migration share are associated with a large *inrease* in social care spending by population aged 65 and above ($p<0.01$). The estimated coefficient shows that a one percentage point increase in the AC-12 migrant shock measure leads to a GBP 56 increase in social care spending by the population aged 65 and above.

We next turn to the local preferences for redistribution channel in column (5) and (6) of table 5.12. The association of AC-12 migration inflows with the share of local council seats held by Conservatives is negative and not significantly different from zero (6). In fact, we find suggestive evidence that it is the share in Labour held seats that increased following AC-12 migration inflows ($p<0.1$; column 5). Our data does not allow us to disentangle whether the shift towards Labour votes were in parts caused by AC-12 migrant voters themselves or if their presence causally affected the voting behaviour of the local non-migrant population. However, our results allow us to conclude that, if there was indeed a shift of native preferences towards the less-redistributive Conservative party, this shift was smaller than any excess Labour votes cast by migrants themselves, such that the aggregate local preferences showed no signs of a preference shift towards less redistribution as a response to more heterogeneity. In appendix 5.E, table 5.18, we find suggestive evidence that AC-12 did indeed increase UKIP votes by approximately the same magnitude as Conservative seats decreased, although these results are only stable when including controls.

Finally, columns (7) and (8) show the association of AC-12 migration with net total

UK-internal migrant outflows and the net internal outflow of UK-borns respectively. Unlike previous research by Sá [2015] and Giuntella et al. [2018], who analyse migrants as a homogenous group, we find no evidence for a "native flight" following the inflow of AC-12 migrants into local authority areas. Instead, our estimates in fact suggest a decline in the net outflow of both total and UK native-born internal migrants in response to AC-12 migration.

Taken together, we interpret these results as strongly suggestive of a greater role played by the demographic channel. While the results on native flight and the increase in the more pro-redistributive platform go against the preferences channel, the fact AC-12 migrants were on average younger and more likely to be in employment suggests the demographic channel played a significant role in explaining the reduction in social care spending per capita. Furthermore, education spending per pupil remains stable and can explain the rise in central government transfers per capita for this partially ring-fenced expenditure item. While we cannot fully test this argument, it is also worth noting that this increase in pupils per capita should not be fully attributed to a rise in pupils from AC-12 countries only. As shown in table 5.1, AC-12 migrants were not particularly more likely to have more children than other groups, such that the relative increase in pupils may also reflect the second order effect a reduction in net outflows of potentially young natives with children.

Table 5.13 then turns to a more detailed analysis of the effect of AC-12 migration on local revenue sources by breaking up the local revenue measure into the council tax and the tax base.

	(1) Council tax (Band-D)	(2) Taxable dwellings pc
AC-12 shock	-395.33*	-0.215***
	(235.61)	(0.050)
N	2028	2028
R^2	0.949	0.789
Time FE	Year	Year
Model	DiD	DiD
Full set of controls	Yes	Yes
Sample	LFS/APS	LFS/APS

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Notes: The Council tax (Band-D) is the baseline annual lump-sum Council tax set by local authorities. Taxable dwellings is the number of Band-D equivalent dwellings that are subject to paying Council tax. "pc" refers to per capita. The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population. LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey. The observation period is 2000 to 2015.

Table 5.13: Effect of AC-12 migration on local revenue sources

The results shown in table 5.13 suggest that, while it was indeed the case that the housing stock did not keep up with the increasing number of migrants (column 2), AC-12 migration was also associated with a relative decrease in the local council tax (column 1, p<0.1).

In light of the results on local voting patterns and internal migration responses to AC-12 inflows shown in columns (5) to (8) of table 5.12, we do not interpret these effects as reflecting the importance of the preferences channel. The provision of local housing units naturally lags changes in the local population size, explaining the decline in the tax base per capita as affected local authorities' populations increased drastically. This decline in the number of taxable dwellings relative to the local population did likely not require a compensation by increasing the local council tax. This interpretation is corroborated by the fact that the reduction in social care spending per capita was not conducted at the expense, but rather at the benefit of the most vulnerable populations (column 4, table 5.12). Indeed, the most plausible interpretation of our findings is that the inflow of a dynamic and young population eased pressure on local authorities who could now spend relatively more on the social care of those aged 65 and above.

In summary, our analysis of mechanisms that explain the shifts in local authority expenditure and revenue associated with AC-12 migration inflows leads us to three main conclusions: First, the changes AC-12 migrants caused to local authority expenditure and revenue patterns were in large parts due to the shifts these migrants caused to lo-

cal demographics and the corresponding institutional responses that were triggered by the resulting changes in local service demand. Second, AC-12 inflows were significantly associated with a relative increase in internal migration. The increasing share of pupils associated with AC-12 migration was likely linked to internal migrant inflows into the same local authorities leading to increased education spending per capita. Finally, the relative drop in demand for local social care services associated with AC-12 migrants did not just result in a large *inrease* in social care spending by the population aged 65 and above, but also allowed affected local authorities to keep their council tax rates relatively low. Overall, our results suggest that the demographic channel was most relevant to explain the impact of AC-12 migration. Thus, they further strengthen previous research findings of an overall net positive fiscal contribution migrants from 2004 EU accession countries made to the UK government budget [Dustmann and Frattini, 2014].

5.8 Conclusion

In this study, we investigated the effect of the large and unexpected wave of Central and Eastern European migrants starting in 2004 on local authority redistributive spending in England. Our results suggest that AC-12 migrants indeed impacted on local authority revenue sources and the local spending mix. We do not find evidence in favour of the hypothesis that these large migration inflows impacted on local preferences for redistributing income. We neither observe voting patterns in local elections that would indicate such a shift in preferences, nor did local residents "vote with their feet" in response to migration inflows. While our results clearly favour shifts in demographics as an explanation for the observed changes in local authority revenue and spending, a limitation of our study is the lack of survey data that could capture preferences for redistribution more precisely.

We interpret our results as a word of caution when relating them to the existing literature. A decrease in public spending can mean a lack of demand from newcomers due to their distinct socio-economic characteristics, rather than the outcome of an increasing local insider-outsider dynamic. Thus, our findings rather lend further support to Dustmann and Frattini [2014] who show that AC-12 migrants were positive net contributors to the UK public purse.

It is worth reflecting on our results in light of the Brexit vote, where anti-immigrant sentiment has played an important role [Meleady et al., 2017, Dennison and Geddes, 2018]. A possible interpretation of our results is that the national-level distaste for immigrants

expressed in the Brexit vote was not driven by local level exposure to foreigners. This explanation finds strong support in a recent study by Becker et al. [2017], who show that local education, income and unemployment levels are strongly correlated with the local Vote Leave share, whereas exposure to EU migrants has little explanatory power. Such an interpretation would further corroborate the necessity to conduct studies relating the presence of migrants to preferences for redistribution on the subnational rather than the national level if the aim is to study direct exposure: On the national level, changes in preferences for redistribution may capture an increased fear of foreigners in areas not necessarily exposed to migration.

Finally, the interpretation of our findings with regards to the sustainability of social care services in England requires a careful reflection. On the one hand, we show that the distinct demographics of AC-12 migrants eased the pressure these means-tested services face in England in the short-term. On the other hand, those migrants who arrived as part of the post-accession waves are increasingly getting older and will likely demand social care services in larger numbers in the future. Using migration as a tool to permanently ease the pressure on local service provision would thus require a continuous inflow of migrants. However, migration inflows from EU countries to the UK have slowed down in response to recent political developments in the UK, with net migration flows from the 2004 Central and Eastern European countries turning negative in 2018 for the first time [Sumption and Vargas-Silva, 2021]. These developments are likely to have repercussions for local authority revenue and spending in the near future.

Appendix

5.A Main results with controls - full table

Table 5.14 shows the coefficients of the AC-12 shock as well as the coefficients of all our control variables on our main outcomes of interest.

Table 5.14: Main results: All controls

	(1) Total expenditure pc	(2) Central government transfers pc	(3) Locally raised budget pc	(4) Social care pc	(5) Education pc	(6) Other pc
AC-12 shock	526.82	978.15**	-451.32***	-414.47**	1179.50**	-238.21
	(457.81)	(472.17)	(87.03)	(204.35)	(564.80)	(152.78)
Share EU-15	1291.99**	1482.37***	-190.38	33.77	1062.35***	195.87
	(514.40)	(441.90)	(156.07)	(230.67)	(365.07)	(231.50)
Share non-EU	247.52	406.58	-159.06**	-62.79	265.37	44.94
	(251.84)	(277.55)	(67.62)	(109.33)	(242.30)	(98.83)
Unemployment rate	51.85	75.64	-23.78	649.06***	-423.88	-173.33
	(343.96)	(355.50)	(106.79)	(187.60)	(405.00)	(145.36)
Population	-0.0020***	-0.0018***	-0.0003**	-0.0006***	-0.0008***	-0.0006***
	(0.0003)	(0.0003)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
N	2028	2028	2028	2028	2028	2028
R^2	0.890	0.834	0.826	0.790	0.757	0.654
Time FE	Year	Year	Year	Year	Year	Year
Model	DiD	DiD	DiD	DiD	DiD	DiD
Sample	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS	LFS/APS

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Notes: Outcomes expressed in per capita terms (pc). The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey. The observation period is 2000 to 2015.

5.B Local authority spending and funding

Table 5.15 is based on Sandford [2018, p.19] and provides a comprehensive overview of how the different functions within two-tier authorities are split between the upper tier (the county) and the lower tier (the district).

Function	Tier
Births, deaths and marriage registration	County
Children's services	County
Concessionary travel	County
Consumer protection	County
Education	County
Emergency planning	County
Highways, street lighting and traffic management	County
Libraries	County
Minerals and waste planning	County
Passenger transport (buses) and transport planning	County
Public health	County
Social care services	County
Trading standards	County
Waste disposal	County
Building regulations	District
Burials and cremations	District
Coastal protection	District
Community safety	District
Council tax and business rates	District
Elections and electoral registration	District
Environmental health	District
Housing	District
Licensing	District
Markets and fairs	District
Public conveniences	District
Sports centres, parks, playing fields	District
Street cleaning	District
Waste collection and recycling	District
Arts and recreation	County and District
Economic development	County and District
Museums and galleries	County and District
Parking County	County and District
Planning	County and District
Tourism	County and District

Table 5.15: Overview of local authority services in England by government tier

Table 5.16 provides an overview of social care services in English local authorities. [8]

Service	Expenditure items included	Means testing (as of 2018)
Adult social care	Physical support Sensory support Memory and cognition support Learning disability support Mental health support Social support (Substance abusers and asylum seekers) Assistive equipment Social care activities Early intervention counselling	Needs test: Carried out by local council. Asset test: Lower asset threshold of £14,250; Upper asset threshold £23,250. Full coverage below lower threshold, shared at the discretion of local authorities in between. Income test: Entitled to £24.90/week of Personal Expense Allowance for care home residents; Minimum Income Guarantee of between £91.40/week and £144.30/week for other settings.
Child social care	Children centres Children looked after Family support services Other children and family support Youth justice Safeguarding children's safety Services for young people	Needs test: Carried out by local authorities within 45 days after referral.

Notes: Own table based on Chartered Institute for Public Finances and Account data; [Jarrett, 2017] ; NHS[9]

Table 5.16: Overview of social care services in England

[8]Note that the thresholds for means testing changed over our observation period; however the general setting remains the same

5.C Spatial distribution of other migrant groups

Figure 5.7 shows the change in the shares within the population of non-EU migrants and EU-15 migrants between 2004, the year of enlargement, and 2015.

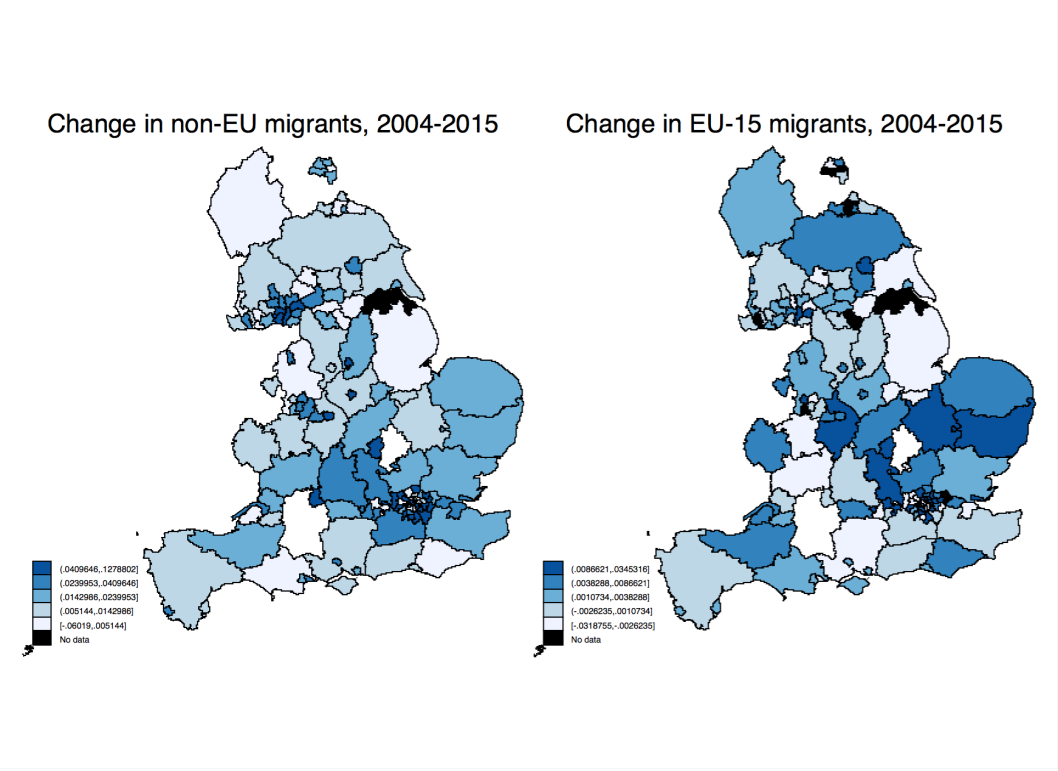


Figure 5.7: Changes in non-EU and EU-15 migrants by English local authority

5.D 2001 Census variables for matching

Table 5.17 shows the full list of 2001 local authority characteristics drawn from the 2001 UK Census. The variables highlighted in grey were selected as best predictors for AC-12 inflows between 2004 and 2015.

Variable	Mean	Std. Dev.	N
Tax base	110.802	94.351	143
Council tax per capita required	304.85	446.339	143
Central government transfer per capita	839.677	1221.178	143
Share spent on social care	0.235	0.048	143
Share spent on education	0.525	0.066	143
Share spent on other items	0.24	0.068	143
Unemployment	0.059	0.025	143
Population	327418.182	257845.906	143
Household share - not deprived	0.413	0.061	143
Household share - deprived in one dimension	0.329	0.019	143
Household share - deprived in two dimensions	0.196	0.036	143
Household share - deprived in three dimensions	0.056	0.019	143
Household share - deprived in all dimensions	0.006	0.003	143
Household share - owns house	0.686	0.123	143
Household share - socially rented	0.196	0.093	143
Household share - socially rented from Council	0.137	0.08	143
Household share - privately rented	0.101	0.053	143
Share houses unoccupied in local authority	0.039	0.025	143
Share houses unoccupied - secondary houses	0.007	0.023	143
Share houses unoccupied - vacant	0.032	0.012	143
Share population in rural area	17.644	24.566	142
Share EU-15 born	0.015	0.015	143
Share AC-12 born	0.006	0.008	143
Share non-EU born	0.079	0.083	143
Share industry: Agriculture	0.011	0.012	143
Share industry: Fishing	0.001	0.001	143
Share industry: Mining	0.002	0.002	143
Share industry: Manufacturing	0.144	0.055	143
Share industry: Electricity	0.007	0.004	143
Share industry: Construction	0.065	0.016	143
Share industry: Whole Sale	0.167	0.025	143
Share industry: Hotels	0.05	0.024	143
Share industry: Transport	0.073	0.019	143
Share industry: Financial	0.051	0.029	143
Share industry: Real Estate	0.137	0.056	143
Share industry: Public sector	0.055	0.016	143
Share industry: Education	0.076	0.012	143
Share industry: Health	0.107	0.018	143
Share industry: Domestic work	0.001	0.001	143
Share aged 64+	0.072	0.016	143

Table 5.17: Summary statistics matching variables

5.E UKIP results

Table 5.18 shows the impact of our AC-12 shock measure on the electoral results of UKIP in local elections that took place between 2000 and 2015.

	(1)	(2)
	Vote share UKIP	Vote share UKIP
AC-12 shock	0.78	0.37*
	(0.23)	(0.22)
N	878	774
R^2	0.715	0.752
Time FE	Year	Year
Model	DiD	DiD
Full set of controls	No	Yes
Sample	LFS/APS	LFS/APS

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Notes: Notes: Outcomes expressed in percentage points. The AC-12 shock is defined as the difference in AC-12 population shares in a given local authority-year and its 2001 baseline share as defined in equation 5.1 of section 5.4.1. All regressions include local authority fixed effects. The full set of controls refers to the share of EU-15 migrants, the share of non-EU migrants, the local authority unemployment rate and the total local population. LFS/APS refers to the UK Labour Force Survey and its boost samples in the Annual Population Survey. The observation period is 2000 to 2015.

Table 5.18: Effect of AC-12 migration on UKIP vote shares

Bibliography

G. J. Abel, M. Brottrager, J. C. Cuaresma, and R. Muttarak. Climate, Conflict and Forced Migration. *Global Environmental Change*, 54:239–249, 2019. doi: 10.1016/j.gloenvcha.2018.12.003.

A. F. Abraido-Lanza, B. P. Dohrenwend, D. S. Ng-Mak, and J. B. Turner. The latino mortality paradox: a test of the" salmon bias" and healthy migrant hypotheses. *American journal of public health*, 89(10):1543–1548, 1999.

M. Akbulut-Yuksel. Children of war the long-run effects of large-scale physical destruction and warfare on children. *Journal of Human resources*, 49(3):634–662, 2014.

C. G. Aksoy and P. Poutvaara. Refugees' self-selection into europe: who migrates where? *Available at SSRN 3373580*, 2019.

C. G. Aksoy, P. Poutvaara, and F. Schikora. First time around: Local conditions and multi-dimensional integration of refugees. *IZA DP No. 13914*, 2020.

M. Albahari. From right to permission: Asylum, mediterranean migrations, and europe's war on smuggling. *Journal on Migration and Human Security*, 6(2):121–130, 2018.

M. Aleksynska and A. Tritah. Occupation–education mismatch of immigrant workers in europe: Context and policies. *Economics of Education Review*, 36:229–244, 2013.

Y. Alem and M. Söderbom. Household-level consumption in urban ethiopia: the effects of a large food price shock. *World development*, 40(1):146–162, 2012.

A. Alesina, E. Glaeser, and B. Sacerdote. Why doesn't the us have a european-style welfare system? Technical report, National bureau of economic research, 2001.

A. Alesina, E. Glaeser, and E. L. Glaeser. *Fighting poverty in the US and Europe: A world of difference*. Oxford University Press, 2004.

A. Alesina, E. Murard, and H. Rapoport. Immigration and preferences for redistribution in europe. Technical report, National Bureau of Economic Research, 2019.

J. G. Altonji and D. Card. *The effects of immigration on the labor market outcomes of less-skilled natives*. Routledge, 2018.

S. Andersen, G. W. Harrison, M. I. Lau, and E. E. Rutstroem. Eliciting risk and time preferences. *Econometrica, vol.76(3)*, 2008.

B. Anderson, M. Ruhs, B. Rogaly, and S. Spencer. Fair enough. *Central and East European migrants in low-wage employment in the UK. Joseph Rowntree Foundation*, 2006.

J. D. Angrist and J.-S. Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2008.

J. Arsenijević, E. Schillberg, A. Ponthieu, L. Malvisi, W. A. E. Ahmed, S. Argenziano, F. Zamatto, S. Burroughs, N. Severy, C. Hebting, et al. A crisis of protection and safe passage: violence experienced by migrants/refugees travelling along the western balkan corridor to northern europe. *Conflict and health*, 11(1):1–9, 2017.

J. Arsenijević, D. Burtscher, A. Ponthieu, N. Severy, A. Contenta, S. Moissaing, S. Argenziano, F. Zamatto, R. Zachariah, E. Ali, et al. "i feel like i am less than other people": Health-related vulnerabilities of male migrants travelling alone on their journey to europe. *Social Science & Medicine*, 209:86–94, 2018.

K. Ash and N. Obradovich. Climatic Stress, Internal Migration, and Syrian Civil War Onset. *Journal of Conflict Resolution*, 64(1):3–31, 2020. doi: 10.1177/0022002719864140.

M. Auffhammer, S. M. Hsiang, W. Schlenker, and A. Sobel. Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy*, 7(2):181–198, 2013. doi: 10.1093/reep/ret016.

M. Barslund, M. Busse, K. Lenaerts, L. Ludolph, and V. Renman. Integration of refugees: lessons from bosnians in five eu countries. *Intereconomics*, 52(5):257–263, 2017a.

M. Barslund, A. Di Bartolomeo, L. Ludolph, et al. Gender inequality and integration of non-eu migrants in the eu. *CEPS Policy Insights, CEPS, Brussels*, 2017b.

T. J. Bartik. Who benefits from state and local economic development policies? 1991.

M. Battisti, G. Peri, and A. Romiti. Dynamic effects of co-ethnic networks on immigrants' economic success. Technical report, National Bureau of Economic Research, 2016.

S. Bazzi. Wealth Heterogeneity and the Income Elasticity of Migration. *American Economic Journal: Applied Economics*, 9(2):219–55, 2017. doi: https://doi.org/10.1257/app.20150548.

S. Bazzi and C. Blattman. Economic Shocks and Conflict: Evidence from Commodity Prices. *American Economic Journal: Macroeconomics*, 6(4):1–38, 2014. doi: http://dx.doi.org/10.1257/mac.6.4.1.

S. O. Becker and A. Ferrara. Consequences of forced migration: A survey of recent findings. *Labour Economics*, 59:1–16, 2019.

S. O. Becker, T. Fetzer, et al. Does migration cause extreme voting? *Center for Competitive Advantage in the Global Economy and The Economic & Social Research Council*, pages 1–54, 2016.

S. O. Becker, T. Fetzer, and D. Novy. Who voted for brexit? a comprehensive district-level analysis. *Economic Policy*, 32(92):601–650, 2017.

S. O. Becker, T. Fetzer, et al. *Has Eastern European Migration Impacted UK-born Workers?* University of Warwick, Department of Economics, 2018.

M. Beine, F. Docquier, and Ç. Özden. Diasporas. *Journal of Development Economics*, 95 (1):30–41, 2011.

M. Beine, G. Charness, A. Dupuy, and M. Joxhe. Shaking things up: On the stability of risk and time preferences. *IZA Discussion Papers 13084*, 2020.

M. Beiser. Changing time perspective and mental health among southeast asian refugees. *Culture, Medicine and Psychiatry*, 11(4):437–464, 1987.

B. Bell, F. Fasani, and S. Machin. Crime and immigration: Evidence from large immigrant waves. *Review of Economics and statistics*, 21(3):1278–1290, 2013.

M. F. Bellemare. Rising Food Prices, Food Price Volatility, and Social Unrest. *American Journal of Agricultural Economics*, 97(1):1–21, 2015. doi: 10.1093/ajae/aau038.

M. F. Bellemare, J. Fajardo-Gonzalez, and S. R. Gitter. Foods and Fads: The Welfare Impacts of Rising Quinoa Prices in Peru. *World Development*, 112:163–179, 2018. doi: https://doi.org/10.1016/j.worlddev.2018.07.012.

A. Belloni, V. Chernozhukov, and C. Hansen. Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies, 81, 608-650*, 2014.

A. Bendl. *AsylwerberInnen in Niederösterreich–Integration (nicht) erwünscht?* PhD thesis, uniwien, 2014.

J. Berazneva and D. R. Lee. Explaining the African Food Riots of 2007–2008: An Smpirical Analysis. *Food Policy*, 39:28–39, 2013. doi: 10.1016/j.foodpol.2012.12.007.

M. Berlemann and M. F. Steinhardt. Climate Change, Natural Disasters, and Migration—a Survey of the Empirical Evidence. *CESifo Economic Studies*, 63(4):353–385, 2017. doi: 10.1093/cesifo/ifx019.

E. Berman, K. Lang, and E. Siniver. Language-skill complementarity: returns to immigrant language acquisition. *Labour Economics*, 10(3):265–290, 2003.

N. Berman and M. Couttenier. External Shocks, Internal Shots: the Geography of Civil Conflicts. *Review of Economics and Statistics*, 97(4):758–776, 2015. doi: 10.1162/REST_a_00521.

N. Berman, M. Couttenier, D. Rohner, and M. Thoenig. This Mine is Mine! How Minerals Fuel Conflicts in Africa. *American Economic Review*, 107(6):1564–1610, 2017. doi: 10.1257/aer.20150774.

M. Bertrand, E. Duflo, and S. Mullainathan. How much should we trust differences-indifferences estimates? *The Quarterly journal of economics*, 119(1):249–275, 2004.

T. J. Besley and T. Persson. The Incidence of Civil War: Theory and Evidence. Technical report, National Bureau of Economic Research, 2008. NBER Working Paper.

A. Bindler and N. Ketel. Scaring or scarring? labour market effects of criminal victimisation. *ECONtribute Discussion Papers Series 030*, 2019.

R. Black, S. R. Bennett, S. M. Thomas, and J. R. Beddington. Migration as Adaptation. *Nature*, 478(7370):447–449, 2011. doi: 10.1038/478477a.

C. Blattman and J. Annan. The consequences of child soldiering. *The review of economics and statistics*, 92(4):882–898, 2010.

H. Bleakley and A. Chin. Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and statistics*, 86(2):481–496, 2004.

A. Bloch. Refugees in the uk labour market: The conflict between economic integration and policy-led labour market restriction. *Journal of Social Policy*, 37(1):21–36, 2008.

S. Bonfanti and T. Xenogiani. Migrants' skills: Use, mismatch and labour market outcomes–a first exploration of the international survey of adult skills (piaac). 2014.

G. J. Borjas. The economic analysis of immigration. *Handbook of labor economics*, 3: 1697–1760, 1999.

V. Bosetti, C. Cattaneo, and G. Peri. Should They Stay or Should They Go? Climate Migrants and Local Conflicts. *Journal of Economic Geography*, 2020. doi: 10.1093/jeg/lbaa002.

L. P. Boustan. Was postwar suburbanization ?white flight?? evidence from the black migration. *The Quarterly Journal of Economics*, 125(1):417–443, 2010.

C. Brell, C. Dustmann, and I. Preston. The labor market integration of refugee migrants in high-income countries. *Journal of Economic Perspectives*, 34(1):94–121, 2020.

N. Breslau, E. L. Peterson, and L. R. Schultz. A second look at prior trauma and the post-traumatic stress disorder effects of subsequent trauma: a prospective epidemiological study. *Archives of General Psychiatry*, 65(4):431–437, 2008.

R. Brown, M. Montalva, D. Thomas, and A. Velásquez. Impact of violent crime on risk aversion: Evidence from the mexican drug war. *The Review of Economics and Statistics, MIT Press, vol. 101(5)*, 2019.

S. Brown, J. Roberts, and K. Taylor. Reservation wages, labour market participation and health. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173 (3):501–529, 2010.

H. Brücker, P. Jaschke, S. Keita, and R. Konle-Seidl. Zum gesetzentwurf der bundesregierung über duldung bei ausbildung und beschäftigung sowie zu den anträgen der fraktionen der fdp, bündnis 90/die grünen und die linke. Technical report, IAB-Stellungnahme, 2019.

M. Brückner and A. Ciccone. International Commodity Prices, Growth and the Outbreak of Civil War in Sub-Saharan Africa. *The Economic Journal*, 120(544):519–534, 2010. doi: 10.1111/j.l468-0297.2010.02353x.

M. Burke and K. Emerick. Adaptation to Climate Change: Evidence from US Agriculture. *American Economic Journal: Economic Policy*, 8(3):106–40, 2016.

M. Burke, S. M. Hsiang, and E. Miguel. Global Non-linear Effect of Temperature on Economic Production. *Nature*, 527(7577):235–239, 2015.

M. B. Burke, E. Miguel, S. Satyanath, J. A. Dykema, and D. B. Lobell. Warming Increases the Risk of Civil War in Africa. *Proceedings of the National Academy of Sciences*, 106 (49):20670–20674, 2009. doi: 10.1073ˆpnas.090799810.

B. C. Cadena and B. J. Keys. Human capital and the lifetime costs of impatience. *American Economic Journal: Economic Policy, 7(3)*, 2015.

M. Callen. Catastrophes and time preference: Evidence from the indian ocean earthquake. *Journal of Economic Behavior & Organization, Elsevier, vol. 118*, 2015.

M. Callen, M. Isaqzadeh, J. D. Long, and C. Sprenger. Violence and risk preference: Experimental evidence from afghanistan. *American Economic Review*, 104(1):123–48, 2014.

L. Cameron and M. Shah. Risk-taking behavior in the wake of natural disasters. *Journal of Human Resources, vol. 50(2)*, 2015.

A. Cangiano. Uk data sources on international migration and the migrant population: a review and appraisal. *COMPAS Research Resource Paper*, 2010.

D. Card. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64, 2001.

D. Card, C. Dustmann, and I. Preston. Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, 10(1):78–119, 2012.

T. A. Carleton and S. M. Hsiang. Social and Economic Impacts of Climate. *Science*, 353 (6304):aad9837, 2016. doi: 10.1126/science.aad98.

E. U. Cascio and E. G. Lewis. Cracks in the melting pot: immigration, school choice, and segregation. *American Economic Journal: Economic Policy*, 4(3):91–117, 2012.

A. Cassar, A. Healy, and C. von Kessler. Trust, risk, andtime preferences after a natural disaster: Experimental evidence from thailand. *World Development, vol. 94, 90-105*, 2017.

C. Cattaneo and G. Peri. The Migration Response to Increasing Temperatures. *Journal of Development Economics*, 122:127–146, 2016. doi: 10.1016/j.jdeveco.2016.05.004.

C. Cattaneo, M. Beine, C. J. Fröhlich, D. Kniveton, I. Martinez-Zarzoso, M. Mastrorillo, K. Millock, E. Piguet, and B. Schraven. Human Migration in the Era of Climate Change. *Review of Environmental Economics and Policy*, 13(2):189–206, 2019.

J. E. Cavanaugh and A. A. Neath. The akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics*, 11(3):e1460, 2019.

D. Chetverikov, Z. Liao, and V. Chernozhukov. On cross-validated lasso. *ArXiv Working Paper No. arXiv:1605.02214*, 2019.

B. R. Chiswick. Speaking, reading, and earnings among low-skilled immigrants. *Journal of labor economics*, 9(2):149–170, 1991.

B. R. Chiswick and P. W. Miller. The endogeneity between language and earnings: International analyses. *Journal of labor economics*, 13(2):246–288, 1995.

B. R. Chiswick and P. W. Miller. The international transferability of immigrants' human capital. *Economics of Education Review*, 28(2):162–169, 2009.

M. A. Clemens. Economics and emigration: Trillion-dollar bills on the sidewalk? *Journal of Economic perspectives*, 25(3):83–106, 2011.

M. A. Clemens. Does Development Reduce Migration? In *International Handbook on Migration and Economic Development*. Edward Elgar Publishing, 2014.

E. Commission. Mapping of refugee integration policies across eu member states. 2016a.

E. Commission. Evaluation of the application of the recast qualification directive (2011/95/eu). 2016b.

P. Connor. Explaining the refugee gap: Economic outcomes of refugees versus other immigrants. *Journal of Refugee Studies*, 23(3):377–397, 2010.

Copernicus Climate Change Service (C3S). ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate. Technical report, 2017. URL https://cds.climate.copernicus.eu/cdsapp#!/home.

K. E. Cortes. Are refugees different from economic immigrants? some empirical evidence on the heterogeneity of immigrant groups in the united states. *Review of Economics and Statistics*, 86(2):465–480, 2004.

M. Couttenier, V. Petrencu, D. Rohner, and M. Thoenig. The violent legacy of conflict: Evidence on asylum seekers, crime, and public policy in switzerland. *American Economic Review*, 109(12):4378–4425, 2019.

D. Crosier and D. Kocanova. *Integrating Asylum Seekers and Refugees into Higher Education in Europe: National Policies and Measures. Eurydice Report.* ERIC, 2019.

M. Dahlberg, K. Edmark, and H. Lundqvist. Ethnic diversity and preferences for redistribution. *Journal of Political Economy*, 120(1):41–76, 2012.

A. Damas de Matos and T. Liebig. The qualifications of immigrants and their value in the labour market: A comparison of europe and the united states. *Matching economic migration with labour market needs*, pages 187–228, 2014.

A. P. Damm. Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, 27(2):281–314, 2009.

T. H. Dao, F. Docquier, C. Parsons, and G. Peri. Migration and development: Dissecting the anatomy of the mobility transition. *Journal of Development Economics*, 132:88–101, 2018.

R. L. Darville and J. B. Reeves. Social inequality among yugoslav women in directoral positions. *Sociological Spectrum*, 12(3):279–292, 1992.

F. D'Auria, K. Mc Morrow, K. Pichelmann, et al. Economic impact of migration flows following the 2004 eu enlargement process-a model based analysis. Technical report, Directorate General Economic and Financial Affairs (DG ECFIN), European . . . , 2008.

J. De Winne and G. Peersman. The Impact of Food Prices on Conflict Revisited. *Journal of Business & Economic Statistics*, pages 1–14, 2019. doi: 10.1080/07350015.2019.1684301.

A. Deaton. Household Survey Data and Pricing Policies in Developing Countries. *The World Bank Economic Review*, 3(2):183–210, 1989.

S. Decker and H. Schmitz. Health shocks and risk aversion. *Journal of Health Economics, Elsevier, vol. 50(C)*, 2016.

M. Dell, B. F. Jones, and B. A. Olken. What Do We Learn From the Weather? The New Climate–Economy Literature. *Journal of Economic Literature*, 52(3):740–798, 2014. doi: 10.1257/jel.52.3.740.

S. DellaVigna and M. D. Paserman. Job search and impatience. *Journal of Labor Economics, vol.23(3)*, 2005.

M. Demeke, G. Pangrazio, and M. Maetz. Country Responses to the Food Security Crisis: Nature and Preliminary Implications of the Policies Pursued. 2009. FAO report.

J. Dennison and A. Geddes. Brexit and the perils of 'europeanised'migration. *Journal of European Public Policy*, 25(8):1137–1153, 2018.

E. Dinas, K. Matakos, D. Xefteris, and D. Hangartner. Waking up the golden dawn: does exposure to the refugee crisis increase support for extreme-right parties? *Political analysis*, 27(2):244–254, 2019.

F. Docquier and Z. Iftikhar. Brain drain, informality and inequality: a search-and-matching model for sub-saharan africa. *Journal of International Economics*, 120:109–125, 2019.

F. Docquier and H. Rapoport. Globalization, brain drain, and development. *Journal of economic literature*, 50(3):681–730, 2012.

P. Dolan, G. Loomes, T. Peasgood, and A. Tsuchiya. Estimating the intangible victim costs of violent crime. *British Journal of Criminology*, 45(6):958–976, 2005.

A. Dorward. The Short- and Medium-Term Impacts of Rises in Staple Food Prices. *Food Security*, 4(4):633–645, 2012. doi: https://doi.org/10.1007/s12571-012-0210-3.

S. Drinkwater, J. Eade, and M. Garapich. Poles apart? eu enlargement and the labour market outcomes of immigrants in the united kingdom. *International Migration*, 47(1): 161–190, 2009.

O. Dube and J. F. Vargas. Commodity Price Shocks and Civil Conflict: Evidence from Colombia. *The Review of Economic Studies*, 80(4):1384–1421, 2013. doi: 10.1093/restud/rdt009.

E. Duflo. Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *American Economic Review*, 91(4):795–813, 2001. doi: 10.1257/aer.91.4.795.

E. Duflo. The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia. *Journal of Development Economics*, 74(1): 163–197, 2004. doi: https://doi.org/10.1016/j.jdeveco.2003.12.008.

H. O. Duleep and M. C. Regets. Immigrants and human-capital investment. *American Economic Review*, 89(2):186–191, 1999.

J.-C. Dumont and O. Monso. Matching educational background and employment: a challenge for immigrants in host countries. *International Migration Outlook*, pages 131–159, 2007.

J.-C. Dumont, T. Liebig, J. Peschner, F. Tanay, and T. Xenogiani. How are refugees faring on the labour market in europe? a first evaluation based on the 2014 eu labour force survey ad hoc module. *European Commission DG Employment Working Paper*, 1, 2016.

C. Dustmann. Differences in the labor market behavior between temporary and permanent migrant women. *Labour Economics*, 4(1):29–46, 1997.

C. Dustmann. Temporary migration, human capital, and language fluency of migrants. *scandinavian Journal of Economics*, 101(2):297–314, 1999.

C. Dustmann and F. Fabbri. Language proficiency and labour market performance of immigrants in the uk. *The economic journal*, 113(489):695–717, 2003.

C. Dustmann and T. Frattini. The fiscal effects of immigration to the uk. *The economic journal*, 124(580):F593–F643, 2014.

C. Dustmann and A. Okatenko. Out-migration, Wealth Constraints, and the Quality of Local Amenities. *Journal of Development Economics*, 110:52–63, 2014. doi: https://doi.org/10.1016/j.jdeveco.2014.05.008.

C. Dustmann and A. v. Soest. Language fluency and earnings: Estimation with misclassified language indicators. *Review of Economics and Statistics*, 83(4):663–674, 2001.

C. Dustmann and A. Van Soest. Language and the earnings of immigrants. *ILR Review*, 55(3):473–492, 2002.

C. Dustmann, M. Casanova, M. Fertig, I. Preston, and C. M. Schmidt. *The impact of EU enlargement on migration flows*. Number 25/03. Research Development and Statistics Directorate, Home Office, 2003.

C. Dustmann, T. Frattini, and I. P. Preston. The effect of immigration along the distribution of wages. *Review of Economic Studies*, 80(1):145–173, 2013.

C. D. Dziuban and E. C. Shirkey. When is a Correlation Matrix Appropriate for Factor Analysis? Some Decision Rules. *Psychological Bulletin*, 81(6):358, 1974. doi: https://doi.org/10.1037/h0036316.

W. Easterly and R. Levine. Africa's growth tragedy: policies and ethnic divisions. *The quarterly journal of economics*, 112(4):1203–1250, 1997.

C. Eckel, M. El-Gamal, and R. K. Wilson. Risk loving after the storm: A bayesian-network study of hurricane katrina evacuees. *Journal of Economic Behavior & Organization, vol. 69*, 2009.

P.-A. Edin, P. Fredriksson, and O. Åslund. Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment. *The quarterly journal of economics*, 118(1):329–357, 2003.

B. Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.

European Commission. Action plan on integration and inclusion 2021-2027, 2020.

H. Fassmann and U. Reeger. Austria: From guest worker migration to a country of immigration. *IDEA Working Papers*, 1:1–39, 2008.

M. Fazel, R. V. Reed, C. Panter-Brick, and A. Stein. Mental health of displaced and refugee children resettled in high-income countries: risk and protective factors. *The Lancet*, 379(9812):266–282, 2012.

J. Ferwerda. Immigration, voting rights, and redistribution: Evidence from local governments in europe. *The Journal of Politics*, 83(1):321–339, 2021.

T. Fetzer. Did austerity cause brexit? *American Economic Review*, 109(11):3849–86, 2019.

B. Franz. Bosnian refugees and socio-economic realities: Changes in refugee and settlement policies in austria and the united states. *Journal of Ethnic and Migration Studies*, 29(1):5–25, 2003.

B. Franz. *Uprooted and unwanted: Bosnian refugees in Austria and the United States*. Texas A&M University Press, 2005.

M. Freedman, E. Owens, and S. Bohn. Immigration, employment opportunities, and criminal behavior. *American Economic Journal: Economic Policy*, 10(2):117–51, 2018.

G. P. Freeman. Migration and the political economy of the welfare state. *The Annals of the American Academy of Political and Social Science*, 485(1):51–63, 1986.

T. Freitas-Monteiro and L. Ludolph. Barriers to humanitarian migration, victimisation and integration outcomes: evidence from germany. 2021.

L. M. Fuchs, Y. Fan, and C. von Scheve. Value differences between refugees and german citizens: Insights from a representative survey. *International Migration*, 2020.

C. Gathmann and N. Keller. Access to citizenship and the economic assimilation of immigrants. *The Economic Journal*, 128(616):3141–3181, 2018.

A. Gavazza, M. Nardotto, and T. Valletti. Internet and politics: Evidence from uk local elections and local government policies. *The Review of Economic Studies*, 86(5):2092–2135, 2019.

P. J. Georgeoff. *The educational system of Yugoslavia.* US Department of Education, International Education Programs, 1982.

C. Gerdes. The impact of immigration on the size of government: Empirical evidence from danish municipalities. *Scandinavian Journal of Economics*, 113(1):74–92, 2011.

German Federal Office for Migration and Refugees. Press release nr. 002/2020: Widerrufsprüfungen für 2019 fristgerecht erledigt – widerrufsquote bei 3,3 prozent. 2020.

S. Gherghina and D. J. O'Malley. Self-determination during the brexit campaign: Comparing leave and remain messages. *Fédéralisme Régionalisme*, 2019.

L. Giua. Removing the wall around welfare: What do we learn from a8 immigrants in the uk? *Economic Policy*, 35(102):305–356, 2020.

O. Giuntella, C. Nicodemo, and C. Vargas-Silva. The effects of immigration on nhs waiting times. *Journal of health economics*, 58:123–143, 2018.

M. Goodwin and C. Milazzo. Taking back control? investigating the role of immigration in the 2016 vote for brexit. *The British Journal of Politics and International Relations*, 19(3):450–464, 2017.

S. Goosen, A. E. Kunst, K. Stronks, I. E. van Oostrum, D. G. Uitenbroek, and A. J. Kerkhof. Suicide death and hospital-treated suicidal behaviour in asylum seekers in the netherlands: a national registry-based study. *BMC public health*, 11(1):1–8, 2011.

Y. Govind. Is naturalization a passport for better labor market integration? evidence from a quasi-experimental setting. 2021.

C. Gray and E. Wise. Country-specific Effects of Climate Variability on Human Migration. *Climatic Change*, 135(3-4):555–568, 2016.

M. Gray and A. Barford. The depths of the cuts: the uneven geography of local government austerity. *Cambridge journal of regions, economy and society*, 11(3):541–563, 2018.

B. L. Green, M. Korol, M. C. Grace, M. G. Vary, A. C. Leonard, G. C. Gleser, and S. Smitson-Cohen. Children and disaster: Age, gender, and parental effects on ptsd symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry*, 30(6): 945–951, 1991.

E. Greussing. Erfassung ausländischer bildungsabschlüsse im mikrozensus. ergebnisse einer evaluationsbefragung im 3. quartal 2015. *Statistische Nachrichten*, 71(12):958–964, 2016.

J. Grote. The changing influx of asylum seekers in 2014-2016: Responses in germany; focussed study by the german national contact point for the european migration network (emn). *Working Paper BAMF*, 2018.

L. Guichard. Self-selection of asylum seekers: Evidence from germany. *Demography*, pages 1–28, 2020.

L. Guichard, I. Issifou, and S. Keita. Price adjustments on the market for human smuggling: Evidence from a large demand shock. *Working paper*, 2021.

L. Guiso, P. Sapienza, , and L. a. Zingales. Time varying risk aversion. *Journal of Financial Economics, vol. 128(3)*, 2018.

J. Hageboutros. The bosnian refugee crisis: A comparative study of german and austrian reactions and responses. *Swarthmore International Relations Journal*, 1(1):50–60, 2017.

J. Hahn, P. Todd, and W. Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001.

J. Hainmueller, D. Hangartner, and D. Lawrence. When lives are put on hold: Lengthy asylum processes decrease employment among refugees. *Science advances*, 2(8): e1600432, 2016.

J. Hainmueller, D. Hangartner, and D. Ward. The effect of citizenship on the long-term earnings of marginalized immigrants: Quasi-experimental evidence from switzerland. *Science advances*, 5(12):eaay1610, 2019.

M. Halla, A. F. Wagner, and J. Zweimüller. Immigration and voting for the far right. *Journal of the European Economic Association*, 15(6):1341–1385, 2017.

S. Hallegatte and J. Rozenberg. Climate Change Through a Poverty Lens. *Nature Climate Change*, 7(4):250–256, 2017. doi: https://doi.org/10.1038/nclimate3253.

C. Hanaoka, H. Shigeoka, and Y. Watanabe. Do risk preferences change? evidence from the greateast japan earthquake. *American Economic Journal: Applied Economics 10, 298-330*, 2018.

E. A. Hanushek and L. Woessmann. The economics of international differences in educational achievement. In *Handbook of the Economics of Education*, volume 3, pages 89–200. Elsevier, 2011.

N. A. Harmon. Immigration, ethnic diversity, and political outcomes: Evidence from denmark. *The Scandinavian Journal of Economics*, 120(4):1043–1074, 2018.

E. Hauff and P. Vaglum. Integration of vietnamese refugees into the norwegian labor market: the impact of war trauma. *International migration review*, 27(2):388–405, 1993a.

E. Hauff and P. Vaglum. Vietnamese boat refugees: The influence of war and flight traumatization on mental health on arrival in the country of resettlement: A community cohort study of vietnamese refugees in norway. *Acta Psychiatrica Scandinavica*, 88(3): 162–168, 1993b.

K. Hawton. van hk. suicide. *Lancet*, 373(9672):1372–81, 2009.

J. J. Heckman. Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161, 1979.

C. S. Hendrix and S. Haggard. Global Food Prices, Regime Type, and Urban Unrest in the Developing World. *Journal of Peace Research*, 52(2):143–157, 2015. doi: 10.1177/0022343314561599.

K. Hirvonen. Temperature Changes, Household Consumption, and Internal Migration: Evidence from Tanzania. *American Journal of Agricultural Economics*, 98(4):1230–1249, 2016. doi: https://doi.org/10.1093/ajae/aaw042.

R. Hoffmann, A. Dimitrova, R. Muttarak, J. C. Cuaresma, and J. Peisker. A Meta-analysis of Country-level Studies on Environmental Change and Migration. *Nature Climate Change*, 10(10):904–912, 2020. doi: 10.1038/s41558-020-0898-6.

D. J. Hopkins. Politicized places: Explaining where and when immigrants provoke local opposition. *American political science review*, pages 40–60, 2010.

S. M. Hsiang, M. Burke, and E. Miguel. Quantifying the Influence of Climate on Human Conflict. *Science*, 341(6151):1235367, 2013. doi: 10.1126/science.1235367.

C. Hunkler and M. Khourshed. The role of trauma for integration. the case of syrian refugees. *SozW Soziale Welt*, 71(1-2):90–122, 2020.

IOM. *World Migration Report*. International Organization for Migration, 2020.

M. Ivanic, W. Martin, and H. Zaman. Estimating the Short-run Poverty Impacts of the 2010–11 Surge in Food Prices. *World Development*, 40(11):2302–2317, 2012. doi: https://doi.org/10.1016/j.worlddev.2012.03.024.

J. Jacobsen, J. Klikar, and J. Schupp. Scales manual iab-bamf-soep survey of refugees in germany. Technical report, SOEP Survey Papers, 2017.

D. A. Jaeger, J. Ruist, and J. Stuhler. Shift-share instruments and the impact of immigration. Technical report, National Bureau of Economic Research, 2018.

L. Jaitman and S. Machin. Crime and immigration: new evidence from england and wales. *IZA Journal of Migration*, 2(1):1–23, 2013.

P. Jakiela and O. Ozier. The impact of violence on individual risk preferences: Evidence from a natural experiment. *The Review of Economics and Statistics, vol. 101(3)*, 2019.

T. Jarrett. Social care: Government reviews and policy proposals for paying for care since 1997 (england). *House of Commons Library Briefing Paper*, (8000), 2017.

A.-M. Jeannet, T. Heidland, and M. Ruhs. What asylum and refugee policies do europeans want? evidence from a cross-national conjoint experiment. *European union politics*, page 14651165211006838, 2021.

M. Jetter, L. M. Magnusson, and S. Roth. Becoming sensitive: Males' risk and time preferences after the 2008 financial crisis. *European Economic Review, vol. 128(C)*, 2020.

J. Jofre-Monseny, P. Sorribas-Navarro, and J. Vázquez-Grenno. Immigration and local spending in social services: evidence from a massive immigration wave. *International Tax and Public Finance*, 23(6):1004–1029, 2016.

D. W. Johnston, M. A. Shields, and A. Suziedelyte. Victimisation, well-being and compensation: Using panel data to estimate the costs of violent crime. *The Economic Journal*, 128(611):1545–1569, 2018.

M. Kalkuhl and L. Wenz. The Impact of Climate Conditions on Economic Production. Evidence from a Global Panel of Regions. *Journal of Environmental Economics and Management*, 103:102360, 2020. doi: 10.1016/j.jeem.2020.102360.

S. C. Kassenboehmer and J. P. Haisken-DeNew. You're fired! the causal negative effect of entry unemployment on life satisfaction. *The Economic Journal*, 119(536):448–462, 2009.

H. Kauffmann, A. Knapp, E. Novotny, and S. Schoch. Flüchtlinge und schule? erfahrungen. *ZEP: Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik*, 25(2):15–19, 2002.

G. Kemp, J. S. Silva, et al. Partial Effects in Fixed-effects Models. In *United Kingdom Stata Users' Group Meetings 2016*, number 06. Stata Users Group, 2016.

D. Kemptner and S. Tolan. The role of time preferences in educational decision making. *Economics of Education Review, vol. 67(C)*, 2018.

S. Kennedy. Eea nationals: the ?right to reside?requirement for benefits. *House of Commons Library Standard Note (Social Policy Section), SN/SP/5972*, pages 1–20, 2011.

N. Kettlewell. Risk preference dynamics around life events. *Journal of Economic Behavior & Organization, vol. 162*, 2019.

Y. Kitazawa. Hyperbolic Transformation and Average Elasticity in the Framework of the Fixed Effects Logit Model. 2012. doi: 10.4236/tel.2012.22034.

E. Kofman, S. Lukes, A. D'Angelo, and N. Montagna. The equality implications of being a migrant in britain. 2009.

L. Koppensteiner, M. F. & Menezes. Violence and human capital investments. *Journal of Labor Economics, vol. 39(2)*, 2021.

V. Koubi. Climate Change and Conflict. *Annual Review of Political Science*, 22:343–360, 2019. doi: 10.1146/annurev-polisci-050317-070830.

I. Krumpal. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4):2025–2047, 2013.

T. Lancaster. The Incidental Parameter Problem since 1948. *Journal of Econometrics*, 95(2):391–413, 2000. doi: https://doi.org/10.1016/S0304-4076(99)00044-5.

V. Lavy, O. Silva, and F. Weinhardt. The good, the bad, and the average: Evidence on ability peer effects in schools. *Journal of Labor Economics*, 30(2):367–414, 2012.

D. L. Lee, J. McCrary, M. J. Moreira, and J. Porter. Valid t-ratio inference for iv. *arXiv preprint arXiv:2010.05058*, 2020.

G. Leon. Civil conflict and human capital accumulation the long-term effects of political violence in perú. *Journal of Human Resources*, 47(4):991–1022, 2012.

E. Lerner, G. A. Bonanno, E. Keatley, A. Joscelyne, and A. S. Keller. Predictors of suicidal ideation in treatment-seeking survivors of torture. *Psychological trauma: theory, research, practice, and policy*, 8(1):17, 2016.

D. B. Lobell, W. Schlenker, and J. Costa-Roberts. Climate Trends and Global Crop Production since 1980. *Science*, 333:616–620, 2011. doi: 10.1126/science.1204531.

A. Lochmann, H. Rapoport, and B. Speciale. The effect of language training on immigrants' economic integration: Empirical evidence from france. *European Economic Review*, 113:265–296, 2019.

S. Longhi, P. Nijkamp, and J. Poot. A meta-analytic assessment of the effect of immigration on wages. *Journal of economic surveys*, 19(3):451–477, 2005.

S. Maccini and D. Yang. Under the Weather: Health, Schooling, and Economic Consequences of Early-life Rainfall. *American Economic Review*, 99(3):1006–26, 2009. doi: 10.1257/aer.99.3.1006.

S. Mahuteau and R. Zhu. Crime victimisation and subjective well-being: panel evidence from australia. *Health economics*, 25(11):1448–1463, 2016.

U. Malmendier and S. Nagel. Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics, vol. 126(1)*, 2011.

M. Manacorda, A. Manning, and J. Wadsworth. The impact of immigration on the structure of wages: theory and evidence from britain. *Journal of the European economic association*, 10(1):120–151, 2012.

M. Marbach, J. Hainmueller, and D. Hangartner. The long-term impact of employment bans on the economic integration of refugees. *Science Advances*, 4(9):eaap9519, 2018.

L. Marchal and C. Naiditch. How Borrowing Constraints Hinder Migration: Theoretical Insights from a Random Utility Maximization Model. *The Scandinavian Journal of Economics*, 122(2):732–761, 2020. doi: https://doi.org/10.1111/sjoe.12355.

J. McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714, 2008.

E. McGuirk and M. Burke. The Economic Origins of Conflict in Africa. *Journal of Political Economy*, 128(10):3940–3997, 2020. doi: 10.1086/709993.

W. McManus, W. Gould, and F. Welch. Earnings of hispanic men: The role of english language proficiency. *Journal of labor Economics*, 1(2):101–130, 1983.

S. Meier and C. Sprenger. Present-biased preferences and credit card borrowing. *American Economic Journal: Applied Economics vol. 2(1)*, 2010.

S. Meier and C. D. Sprenger. Discounting financial literacy: Time preferences and participation in financial education programs. *Journal of Economic Behavior & Organization, vol. 95*, 2013.

R. Meleady, C. R. Seger, and M. Vermue. Examining the role of positive and negative intergroup contact and anti-immigrant prejudice in b rexit. *British Journal of Social Psychology*, 56(4):799–808, 2017.

K. Millock. Migration and Environment. *Annual Review of Resource Economics*, 7(1): 35–60, 2015. doi: 10.1146/annurev-resource-100814-125031.

N. Minot. *Transmission of World Food Price Changes to Markets in Sub-Saharan Africa.* International Food Policy Research Institute Washington, 2010.

A. Missirian and W. Schlenker. Asylum Applications Respond to Temperature Fluctuations. *Science*, 358(6370):1610–1614, 2017. doi: 10.1126/science.aao0432.

C. Monfreda, N. Ramankutty, and J. A. Foley. Farming the Planet: 2. Geographic Distribution of Crop Areas, Yields, Physiological Types, and Net Primary Production in the Year 2000. *Global Biogeochemical Cycles*, 22(1), 2008.

M. P. Murray. Avoiding invalid instruments and coping with weak instruments. *Journal of economic Perspectives*, 20(4):111–132, 2006.

P. Nijkamp, J. Poot, and M. Sahin. *Migration impact assessment: New horizons*. Edward Elgar Publishing, 2012.

M. Nübling, U. Stößel, H.-M. Hasselhorn, M. Michaelis, and F. Hofmann. Measuring psychological stress and strain at work-evaluation of the copsoq questionnaire in germany. *GMS Psycho-Social Medicine*, 3, 2006.

N. Nunn and N. Qian. US Food Aid and Civil Conflict. *American Economic Review*, 104 (6):1630–66, 2014. doi: 10.1257/aer.104.6.1630.

P. Oreopoulos. Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review*, 96(1):152–175, 2006.

P. Ornstein. The price of violence: Consequences of violent crime in sweden. Technical report, Working Paper, 2017.

E. Oster. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204, 2019.

L. Page, D. A. Savage, and B. Torgler. Variation in risk seeking behaviour following large losses: A natural experiment. *European Economic Review, Elsevier, vol. 71(C)*, 2014.

G. Pascall and N. Manning. Gender and social policy: comparing welfare states in central and eastern europe and the former soviet union. *Journal of European Social Policy*, 10 (3):240–266, 2000.

G. Peri. Immigrants, productivity, and labor markets. *Journal of economic perspectives*, 30(4):3–30, 2016.

T. Pettersson and M. Öberg. Organized Violence, 1989–2019. *Journal of Peace Research*, 57(4):597–613, 2020. doi: 10.1177/0022343320934986.

D. Phillips. Equalisation, incentives and discretion in english local public service provision. *Governing England: Devolution and funding*, page 35, 2018.

M. Porter and N. Haslam. Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: a meta-analysis. *Jama*, 294 (5):602–612, 2005.

C. Raleigh, A. Linke, H. Hegre, and J. Karlsen. Introducing ACLED: an Armed Conflict Location and Event Dataset: Special Data Feature. *Journal of Peace Research*, 47(5): 651–660, 2010. doi: 10.1177/0022343310378914.

C. Raleigh, H. J. Choi, and D. Kniveton. The Devil is in the Details: An Investigation of the Relationships between Conflict, Food Price and Climate across Africa. *Global Environmental Change*, 32:187–199, 2015. doi: 10.1016/j.gloenvcha.2015.03.005.

H. Rapoport and F. Docquier. The economics of migrants' remittances. *Handbook of the economics of giving, altruism and reciprocity*, 2:1135–1198, 2006.

D. K. Ray, J. S. Gerber, G. K. MacDonald, and P. C. West. Climate Variation Explains a Third of Global Crop Yield Variability. *Nature Communications*, 6(1):1–9, 2015. doi: 10.1038/ncomms6989.

E. Reyneri and G. Fullin. Labour market penalties of new immigrants in new and old receiving west european countries. *International Migration*, 49(1):31–57, 2011.

I. Ruiz and C. Vargas-Silva. Differences in labour market outcomes between natives, refugees and other migrants in the uk. *Journal of Economic Geography*, 18(4):855–885, 2018.

F. Sá. Immigration and house prices in the uk. *The Economic Journal*, 125(587):1393–1424, 2015.

M. Sagbakken, I. M. Bregård, and S. Varvin. The past, the present and the future: a qualitative study exploring how refugees' experience of time influences their mental health and well-being. *Frontiers in Sociology*, 5:46, 2020.

M. Sandford. Local government in england: structures. *House Commons Libr London*, 2018.

S. M. Sawyer, P. S. Azzopardi, D. Wickremarathne, and G. C. Patton. The age of adolescence. *The Lancet Child & Adolescent Health*, 2(3):223–228, 2018.

B. Schauberger, S. Archontoulis, A. Arneth, J. Balkovic, P. Ciais, D. Deryng, J. Elliott, C. Folberth, N. Khabarov, C. Müller, T. A. M. Pugh, S. Rolinski, S. Schaphoff,

E. Schmid, X. Wang, W. Schlenker, and K. Frieler. Consistent Negative Response of US Crops to High Temperatures in Observations and Crop Models. *Nature Communications*, 8(1):1–9, 2017. doi: 10.1038/ncomms13931.

W. Schlenker and D. B. Lobell. Robust Negative Impacts of Climate Change on African Agriculture. *Environmental Research Letters*, 5(1):014010, 2010. doi: 10.1088/1748-9326/5/1/014010.

W. Schlenker and M. J. Roberts. Nonlinear Temperature Effects Indicate Severe Damages to US Crop Yields under Climate Change. *Proceedings of the National Academy of Sciences*, 106(37):15594–15598, 2009. doi: 10.1073/pnas.0906865106.

W. Schlenker, W. M. Hanemann, and A. C. Fisher. Water Availability, Degree Days, and the Potential Impact of Climate Change on Irrigated Agriculture in California. *Climatic Change*, 81(1):19–38, 2007. doi: https://doi.org/10.1007/s10584-005-9008-z.

B. Sedova, L. Cizmaziova, and A. Cook. A Meta-Analysis of Climate Migration Literature. 2021. Working Paper.

C. Senik, H. Stichnoth, and K. Van der Straeten. Immigration and natives? attitudes towards the welfare state: evidence from the european social survey. *Social indicators research*, 91(3):345–370, 2009.

O. Shemyakina. The effect of armed conflict on accumulation of schooling: Results from tajikistan. *Journal of Development Economics*, 95(2):186–200, 2011.

P. Simpson. Public spending on adult social care in england. *Institute for Fiscal Studies. https://www. ifs. org. uk/publications/9185*, 2017.

B. Speciale. Does immigration affect public education expenditures? quasi-experimental evidence. *Journal of public economics*, 96(9-10):773–783, 2012.

M. Spence. Job market signalling": The quarterly journal of economics", 1973.

A. Steinmayr. Contact versus exposure: Refugee presence and voting for the far-right. *Review of Economics and Statistics*, pages 1–47, 2020.

G. J. Stigler and G. S. Becker. De gustibus non est disputandum. *The american economic review*, 67(2):76–90, 1977.

A. Sujoldžić, L. Peternel, T. Kulenović, and R. Terzić. Social determinants of health–a comparative study of bosnian adolescents in different cultural contexts. *Collegium antropologicum*, 30(4):703–711, 2006.

M. Sumption and C. Vargas-Silva. Net migration to the uk. *Migration Observatory briefing, COMPAS, University of Oxford*, 2021.

R. Sundberg and E. Melander. Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research*, 50(4):523–532, 2013. doi: 10.1177/0022343313484347.

M. Sutter, M. G. Kocher, D. Rutzler, and S. T. Trautmann. Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review 103 (1): 510-31*, 2013.

M. Tabellini. Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *The Review of Economic Studies*, 87(1):454–486, 2020a.

M. Tabellini. Racial heterogeneity and local government finances: Evidence from the great migration. 2020b.

P. Taylor-Gooby. Why do people stigmatise the poor at a time of rapidly increasing inequality, and what can be done about it? *The political quarterly*, 84(1):31–42, 2013.

R. H. Thaler and S. Benartzi. Save more tomorrow^{TM}: Using behavioral economics to increase employee saving. *Journal of Political Economy vol. 112(S1)*, 2004.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B 58, 267-288*, 1996.

H. Tretter. *Temporary protection für bosnische Flüchtlinge in Europa: Länderberichte*. Verlag Österreich, 2000.

M. Valenta and S. P. Ramet. *The Bosnian diaspora: Integration in transnational communities*. Ashgate Publishing, Ltd., 2011.

M. Valenta and Z. Strabac. The dynamics of bosnian refugee migrations in the 1990s, current migration trends and future prospects. *Refugee Survey Quarterly*, 32(3):1–22, 2013.

M. Velamuri and S. Stillman. The impact of crime victimisation on individual well-being: Evidence from australia. *Labour, Employment and Work in New Zealand*, 2008.

M. Verpoorten, A. Arora, N. Stoop, and J. Swinnen. Self-reported food insecurity in africa during the food price crisis. *Food Policy*, 39:51–63, 2013.

J. von Braun. The Food Crisis Isn't Over. *Nature*, 456(7223):701–701, 2008. doi: doi.org/10.1038/456701a.

M. J. Voors, E. E. Nillesen, P. Verwimp, E. H. Bulte, R. Lensink, and D. P. Van Soest. Violent conflict and behavior: a field experiment in burundi. *American Economic Review*, 102(2):941–64, 2012.

L. Wantchekon, M. Klašnja, and N. Novta. Education and Human Capital Externalities: Evidence from Colonial Benin. *The Quarterly Journal of Economics*, 130(2):703–757, 2015.

R. J. Willis and S. Rosen. Education and self-selection. *Journal of political Economy*, 87 (5, Part 2):S7–S36, 1979.

J. M. Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.

D. Yang. Migrant remittances. *Journal of Economic perspectives*, 25(3):129–52, 2011.

R. Yehuda. Post-traumatic stress disorder. *New England journal of medicine*, 346(2): 108–114, 2002.

D. Ziebarth. Ukip support in local elections: Which factors play a role in determining electoral fortunes? *Available at SSRN 3570053*, 2020.

W. Zwysen. Different patterns of labor market integration by migration motivation in europe: the role of host country human capital. *International Migration Review*, 53(1): 59–89, 2019.